```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    November  8, 2004, 18:41:28 ; Search time 40 Seconds
                                 (without alignments)
                        870.424 Million cell updates/sec

Title:      US-09-940-919-2
Perfect score:  2750
Sequence:   1 MATLLRSKLSNVATSVSNKS........SGFVHSLEGLIEAYRTNAED 525

Scoring table:  BLOSUM62
            Gapop 10.0 , Gapext 0.5

Searched:   478139 seqs, 6618000 residues

Total number of hits satisfying chosen parameters:    478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued Patents_AA:*
          1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
          2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
          3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
          4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
          5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
          6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

          Pred. No. is the number of results predicted by chance to have a
          score greater than or equal to the score of the result being printed,
          and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2700	98.2	525	4	US-09-178-093B-2	Sequence 2, Appli
2	858	31.2	486	4	US-09-178-093B-1	Sequence 1, Appli
3	429	15.6	149	4	US-09-270-767-31957	Sequence 31957, A
4	303	11.0	145	4	US-09-270-767-33686	Sequence 33686, A
5	303	11.0	145	4	US-09-270-767-48903	Sequence 48903, A
6	211.5	7.7	456	4	US-09-976-594-584	Sequence 584, App
7	209	7.6	412	4	US-09-311-021-42	Sequence 42, Appl
8	203.5	7.4	500	4	US-09-178-093B-26	Sequence 26, Appl
9	192.5	7.0	418	4	US-09-640-419C-25	Sequence 25, Appl
10	192	7.0	504	4	US-09-178-093B-28	Sequence 28, Appl
11	183.5	6.7	447	3	US-09-370-253-2	Sequence 2, Appl
12	181.5	6.6	504	4	US-09-919-497-67	Sequence 67, Appl
13	175	6.4	447	3	US-09-370-253-10	Sequence 10, Appl
14	174	6.3	452	4	US-09-640-419C-5	Sequence 5, Appl
15	173	6.3	446	4	US-09-640-419C-26	Sequence 26, Appl
16	164	6.0	447	3	US-09-370-253-6	Sequence 6, Appli
17	156.5	5.7	399	4	US-09-248-796A-20813	Sequence 20813, A
18	155.5	5.7	449	4	US-09-640-419C-24	Sequence 24, Appl
19	154	5.6	432	3	US-09-370-253-12	Sequence 12, Appl
20	153	5.6	462	4	US-09-640-419C-27	Sequence 27, Appl
21	145.5	5.3	483	4	US-09-248-796A-20815	Sequence 20815, A
22	129.5	4.7	511	4	US-09-248-796A-20816	Sequence 20816, A
23	128	4.7	485	3	US-08-362-512A-2	Sequence 2, Appli
24	128	4.7	485	3	US-08-964-939-2	Sequence 2, Appli
25	123.5	4.5	493	1	US-08-362-512A-4	Sequence 4, Appli
26	123.5	4.5	493	4	US-08-964-939-4	Sequence 4, Appli
27	123	4.5	516	4	US-09-328-352-5639	Sequence 5639, Ap

28	121.5	4.4	565	4	US-09-252-991A-33045	Sequence 33045, A
29	117.5	4.3	414	4	US-09-543-681A-4320	Sequence 4320, Ap
30	117	4.3	21	3	US-09-178-093B-10	Sequence 10, Appl
31	116	4.2	259	3	US-09-370-253-4	Sequence 4, Appli
32	115	4.2	473	4	US-09-637-118B-2	Sequence 2, Appli
33	114	4.1	154	4	US-09-270-767-31638	Sequence 31638, A
34	114	4.1	154	4	US-09-270-767-46855	Sequence 46855, A
35	113	4.1	21	4	US-09-178-093B-13	Sequence 13, Appl
36	111.5	4.1	324	4	US-09-328-352-4577	Sequence 4577, Ap
37	111	4.0	458	4	US-09-489-039A-13954	Sequence 13954, A
38	109.5	4.0	159	4	US-09-621-976-5730	Sequence 5730, Ap
39	109.5	4.0	419	4	US-09-252-991A-23245	Sequence 23245, A
40	109	4.0	21	4	US-09-178-093B-5	Sequence 5, Appli
41	108.5	3.9	551	4	US-09-252-991A-20358	Sequence 20358, A
42	108.5	3.9	1040	3	US-09-134-001C-5365	Sequence 5365, Ap
43	108	3.9	21	4	US-09-178-093B-8	Sequence 8, Appli
44	107.5	3.9	512	4	US-09-328-352-6192	Sequence 6192, Ap
45	107.5	3.9	543	4	US-09-107-532A-3717	Sequence 3717, Ap

ALIGNMENTS

```
RESULT 1
US-09-178-093B-2
; Sequence 2, Application US/09178093B
; Patent No. 6660846
; GENERAL INFORMATION:
;   APPLICANT:  Robert H. Edwards
;   APPLICANT:  Richard J. Reimer
;   APPLICANT:  Steve L. McIntire
;   APPLICANT:  Erik M. Jorgenson
;   APPLICANT:  Kim Schuske
;   TITLE OF INVENTION:  Vesicular Amino Acid Transported
;   TITLE OF INVENTION:  Composition and Method
;   FILE REFERENCE:  2002-0005.30
;   CURRENT APPLICATION NUMBER:  US/09/178,093B
;   CURRENT FILING DATE:  2001-08-20
;   PRIOR APPLICATION NUMBER:  60/063,012
;   PRIOR FILING DATE:  1997-10-23
;   NUMBER OF SEQ ID NOS:  50
;   SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH:  525
;   TYPE:  PRT
;   ORGANISM:  Rattus norvegicus
;   FEATURE:
;     NAME/KEY:  CHAIN
;     LOCATION:  (1)...(525)
;   OTHER INFORMATION:  RUNC-47 polypeptide sequence
US-09-178-093B-2

Query Match             98.2%;  Score 2700;  DB 4;  Length 525;
Best Local Similarity   98.5%;  Pred. No. 4e-274;
Matches 518; Conservative 3; Mismatches 3; Indels 2; Gaps 2;

Qy      1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEBAVGFAHCDDLDFEHRQGLQM 60
          |||||||||:||||||||||||||:|||||||||||||||||||||||||||||||||||
Db      1 MATLLRSKLTNVATSVSNKSQAKVSGMFARMGFQAATDEBAVGFAHCDDLDFEHRQGLQM 60

Qy     61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQ-VGGGGEFGGHDKPKITAW 119
          |||||.|||||||||.|||||||||||||:|||||||||||||.|:||||||||||||||
Db     61 DILKSEGEPCGDEGAEPPVEGDIHYQRG-GAPLPPSGSKDQAVGAGEFGGHDKPKITAW 119

Qy    120 EAGWNVTNAIQGMFVLGLPYAILHGGYIGLFLIIFAAVCCYTGKILIACLYEENEDGEV 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 EAGWNVTNAIQGMFVLGLPYAILHGGYIGLFLIIFAAVCCYTGKILIACLYEENEDGEV 179

Qy    180 VRVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVS 239
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 VRVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVS 239

Qy    240 QKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKF 299
```

```
Db  240 QKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKF 299
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  300 YIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAY 359
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  300 YIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAY 359

Qy  360 LTWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFP 419
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  360 LTWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFP 419

Qy  420 ACYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLR 479
        ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  420 ACYGGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLR 479

Qy  480 LLWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  480 LLWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
```

RESULT 2
US-09-178-093B-1
; Sequence 1, Application US/0917809.3B
; Patent No. 6660846
; GENERAL INFORMATION:
; APPLICANT: Robert H. Edwards
; APPLICANT: Richard J. Reimer
; APPLICANT: Steve L. McIntire
; APPLICANT: Erik M. Jorgenson
; APPLICANT: Kim Schuske
; TITLE OF INVENTION: Vesicular Amino Acid Transported
; TITLE OF INVENTION: Composition and Method
; FILE REFERENCE: 2002-0005.30
; CURRENT APPLICATION NUMBER: US/09/178,093B
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/063,012
; PRIOR FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CHAIN
; LOCATION: (1)...(486)
; OTHER INFORMATION: UNC-47 polypeptide sequence
US-09-178-093B-1

```
Query Match     31.2%;  Score 858;  DB 4;  Length 486;
Best Local Similarity  40.4%;  Pred. No. 5.1e-81;
Matches 180;  Conservative  96;  Mismatches 141;  Indels  28;  Gaps  8;

Qy  75 AEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWEAGWNVTNAIQGMFV 134
           :  :       |                                   ||||||||||||
Db  63 SEQPQKDDINKQ-------BEAKDD--GHGE----ASEPISALQAAWNVTNAIQGMFI 107

Qy 135 LGLPYAILHGGYLGLFLLIFAAVVCCTGKILIACLYEENEDGEVVRVRDSYVAIANACC 194
       :|||| | :||| | |    ||  : |||| |||||||                 ||||
Db 108 VGLPIAVKVGGWNSIGAMVGVAYVCWTGVLLIECLYENG-----VKKRKTYREIAD-FY 161

Qy 195 APRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQKSWSIIATAVLLPC 254
         | |   ||   | ||  ||| ||||| |||: : :   :|||       ||| :||:|
Db 162 KPGF----GKWVLAAQLTELLSTCIIYVLVLAADLLQSCFPS--VDKAGWMMITSASLLTC 215

Qy 255 AFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFYIDVKKFPISIGIIV 314
          :  |:  :: ||   : | :|| | | |||| : :         |||||||||:||:|
Db 216 SFLDDLQIVSRLSFFNAISHLIVNLIMVLYCLSFVSQWSFSTITFSLNINTLPTIVGMVV 275

Qy 315 FSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYLTWADETKEVITDNL 374
       | |||:|||| |||||   | | ||::  ||| |: ::      :              |
Db 276 FGYTSHIFLPNLEGNMKKNPAQFNVMLKWSHIAAAVFKVVFGMLGFLITFGELITQEEISNSL 335
US-09-178-093B-1
```

```
Qy 375 PG-SIRAVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGL 433
          ::  :  ::|||||||||||| ||  ||  :: || :|:  ::: | :||::|:: ||
Db 336 PNQSFKILVNLILVVKALLSYPLPFYAVQLLKNLFLGYPQTPFTSCYSPDKSLREWAV 395

Qy 434 TLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLMWRKLLWHQVFFD 493
       |||     | | || | ||| : |||| |:: | : | : ||||| ::  | | ::||| :
Db 396 TLRILVLFTLFVALSVPYLVELMGLVGNITGTMLSFIWPALFHLYIKEKTLNNFEKRFD 455

Qy 494 VAIFVIGGICSVSGFVHSLEGLIEA 518
        :   :|| |: :||:  ||| :
Db 456 QGIIIMGCSVCISGVYFSSMELLRA 480
```

RESULT 3
US-09-270-767-31957
; Sequence 31957, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31957
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-31957

```
Query Match     15.6%;  Score 429;  DB 4;  Length 149;
Best Local Similarity  53.3%;  Pred. No. 8.1e-37;
Matches  81;  Conservative  27;  Mismatches  40;  Indels  4;  Gaps  2;

Qy 124 NVTNAIQGMFVLGLPYAILHGGYLGLFLLIFAAVVCCTGKILIACLYE-ENEDGEVVRV 182
       ||||||||||: ||||| |||||  |  |:  |  :|| :| | | || :|:||||:|||
Db   1 NVTNAIQGMFIVSLPFAVLHGGYVAIVAMVGIAHICCYTGKVLVQCLYEPDPATGQMVRV 60

Qy 183 RDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQKS 242
       |||||||   | | |     |  :  |||::|| |||||||||| :|   |  |  :  :|
Db  61 RDSYVAIAKVCFGPK--LGARAVSIAQLIELLMTCILYVVVCGDLLAGTYPQGSFDSRS 117

Qy 243 WSIIATAVLLPCAFLKNLKRAVSKFSLLCTLAH 274
       |  :  |::|| ||||:|:  |: :| |  | :|
Db 118 WMLFVGIFLLPMGFLKSLRMVSTLSFWCTMSH 149
```

RESULT 4
US-09-270-767-33686
; Sequence 33686, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33686
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-33686

```
Query Match     11.0%;  Score 303;  DB 4;  Length 145;
Best Local Similarity  43.9%;  Pred. No. 1.3e-23;
Matches  61;  Conservative  26;  Mismatches  52;  Indels  0;  Gaps  0;

Qy 381 VVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALV 440
        ||: :| ||:||||||| : |  | : :    :|  || | :|||||:||||:  |  |:
Db   1 MVNFFIVIKAILSYPLPYYAACELIERNFFRGPPKTKFPTIWNLDGELKVWGLGFRVGVI 60
```

```
Qy   441  VFTTLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLWRKLLWHQVFFDVAIFVIG 500
          :||:::|:|:|||::|:|: ||:|||  :::|::  :::||: ::|:    :: :|:
Db    61  VSTILMAIFIPHFSILMGFIGSFTGTMLSFIWPCYFHIKIGHLLDQKEIAKDYLIIGLG 120

Qy   501  GICSVSGFVHSLEGLIEAY 519
          :: |||||  |||||||||
Db   121  VLFGVIGIYDSGNALINAF 139


RESULT 5
US-09-270-767-48903
; Sequence 48903, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48903
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-48903

  Query Match             11.0%;  Score 303; DB 4; Length 145;
  Best Local Similarity   43.9%;  Pred. No. 1.3e-23;
  Matches    61; Conservative  26; Mismatches  52; Indels  0; Gaps  0;

Qy   381  VVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALV 440
          :|||||||||||||::|| |:|:::|  |:::::   |:|::||||:||||||||: : |
Db     1  MVNFPFLVIKALLSYPLPYYAACELLERNFFRGPPKTKPPTIWNLDGELKVWGLGFRVGVI 60

Qy   441  VFTTLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLWRKLLWHQVFFDVAIFVIG 500
          :||:::|:|:|||::|:|: ||:|||  :::|::  :::||: ::|:    :: :|:
Db    61  VSTILMAIFIPHFSILMGFIGSFTGTMLSFIWPCYFHIKIGHLLDQKEIAKDYLIIGLG 120

Qy   501  GICSVSGFVHSLEGLIEAY 519
          :: |||||  |||||||||
Db   121  VLFGVIGIYDSGNALINAF 139


RESULT 6
US-09-976-594-584
; Sequence 584, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 584
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 2466714CD1
US-09-976-594-584

  Query Match              7.7%;  Score 211.5; DB 4; Length 456;
  Best Local Similarity   22.8%;  Pred. No. 2.9e-13;
  Matches   107; Conservative  82; Mismatches 163; Indels 117; Gaps  22;

Qy    73  EGAEAP------VEGDIHVQRGSGAPLPPSGSSKDQVGGGEFGGHDKPKITAWEAGWNVTN 127
          |::||        |  | ||      :|:    |                   |:  |
Db    22  EEAEAEELSPLLSNELHRQRSPGV------------SFG------------LSVFNLMN 56

Qy   128  AIQGMFVLGLPYAILHGGYLGL-FLIIFAAVCCYTGKILIA-CL------YEE---- 173
          ||::|| |:  |  | :::::|  ::: :|:: : :|:   :|::
Db    57  AIMSGGILGLAYVMANTGVFGFSFLLIVALLASYSVHLLSMCIQTAVTSYEDLGLFAF 116

Qy   174  NEDGEVVRVRDSVVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSF 233
          ::||                                   |  |::::|:||::|| || |
Db   117  GLPGKLVVA--GTIIIQNI----------GAMSYLLIKTELPAAIAEFLTGD--YNRY 162

Qy   234  PGLPVSQKSWSIIATAVLLPCAFLKNL------------KAVSKFSLLCTLA 273
          |||||:|||  ::|| |  :||:|:|                :|:| ||||
Db   163  WYLD-GQTLLIICVGIVFPLALLLPKIGFIGYTSSLSFFMMFFALVVIKKWSIPCPL- 220

Qy   274  HFVINILV:AYCLSRARDWAWEKVKFYIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQP 333
          :|::::|  :   ||::|   ::|:|:::|::::||::||:|::|:|||:||::|::::|
Db   221  --TLNYVEKGFQISNVTDC--KPKLFHSKESAYALPTMAFSFLCHTSILPIYCELQSP 276

Qy   334  SEFHCMMWTHIA---ACVLKGLFALVAYLTWADETKEVITDNLPG-----SIRAVV 382
          :|:        |     | :||  |  |||  |:|:|::|:|     |   :|| |
Db   277  SKKR-MQNVTNTAIALSFLIYFISALFGYLTFYD--KVESELLKGYSKYLSHDVVWMTV 332

Qy   383  NIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVF 442
          |:|:|||:::    |::| :|||::  |:   :|  :|    ||: |||||||||:|:
Db   333  KLCILFAVLLTVPLIHFPARKAVTMMFSN-----FPFSW-----IRHFLITL--ALNII 380

Qy   443  TLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLR-----LLWRKL 485
          ::|:|:| |:|:::|||  |:| :||  :|:|:|  ::    :::|
Db   381  IVLLAIYVPDIRNVFGVVGASTSCLIFIFPGLFYLKLSREDFLSWKKL 429


RESULT 7
US-09-311-021-42
; Sequence 42, Application US/09311021
; Patent No. 6706869
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6300-11A
; CURRENT APPLICATION NUMBER: US/09/311,021
; CURRENT FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-311-021-42

  Query Match              7.6%;  Score 209; DB 4; Length 412;
  Best Local Similarity   21.9%;  Pred. No. 4.5e-13;
  Matches    97; Conservative  80; Mismatches 165; Indels 100; Gaps  18;

Qy   123  WNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAV-VCCYTGKILIAC------LYEE 173
          |:::|| |::: ||:::|:||||||:|| :::|||   :|:| : ::|       |||
Db     4  FNLSNAIMGSGILGLAFALANTGIL-LFIVLLTSVTLLSIYSINLLICSKETGCMVYEK 62

Qy   174  NEDGEVVRVRDSVVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSF 233
          ::||                                   :|  |::::|:||::|| || |
Db    63  L--GEQV------------------FGTTGKFVIFGATSLQNTGAMLSYLFIVKNELPSAI 103

Qy   234  PGLPVSQKSWS------------IIATAVLLPCAFLKNLKAV---SKFSLLCTLAHFVINI 279
          |||||:|||               :||| |  |  :||:|      ||:| ||| ||:|
Db   104  KFIMGKEETFPSAWYVDGRVLVVVTFGIILPLCLLKNLGYLGVTSGFSLSC----MVFFL 159

Qy   280  LVIAY------CL-----------SRARDWAWEKVKFYIDV-KKFPISIGIIVFSYTSQI 321
          :::|:      :|           :
```

```
Db   160  IVVIYKKFQIPCIVPELNSTISANSTNADTCTPK---YVTFNSKTVYALPTIAFAFVCHP  216

Qy   322  FLPSLEGNMQQPSEFHCMM--NWTHIAACVLKGLFALVAYLTWADETKEVITDNLPGS--  377

Db   217  SVLPIYSELKDRSQKRMQMVSNISFFAMFVMYFLTAIFGYILTFYDNVQSDLLHKYQSKDD 276

Qy   378  -IRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGLTLR  436

Db   277  ILILTVRLLAVIVAVILTVPVLFF---TVRSSLFELAKKTKFNLCRH-------TVVT    323

Qy   437  CALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLL------WRKLLWHQV  490

Db   324  CILLVINLLVIFFIPSMKDIFGVGVGVTSANMLIFILPSSLYLKITDQDGDKGTQRIWAAL 383

Qy   491  FFDVAIFVIGGICSVSGFVHSL  512

Db   384  FL-------GLGVLSPCQHSL   397
```

RESULT 8
US-09-178-093B-26
; Sequence 26, Application US/09178093B
; Patent No. 6660846
; GENERAL INFORMATION:
; APPLICANT: Robert H. Edwards
; APPLICANT: Richard J. Reimer
; APPLICANT: Steve L. McIntire
; APPLICANT: Erik M. Jorgenson
; APPLICANT: Kim Schuske
; TITLE OF INVENTION: Vesicular Amino Acid Transported
; TITLE OF INVENTION: Composition and Method
; FILE REFERENCE: 2002-0005.30
; CURRENT APPLICATION NUMBER: US/09/178,093B
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/063,012
; PRIOR FILING DATE: 1997-10-23
; SOFTWARE: FastSEQ for Windows Version 3.0
; NUMBER OF SEQ ID NOS: 50
; SEQ ID NO 26
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide (RVT2 polypeptide sequence)
US-09-178-093B-26

Query Match   7.4%;   Score 203.5; DB 4; Length 500;
Best Local Similarity   20.6%; Pred. No. 2.3e-12;
Matches 111; Conservative 92; Mismatches 225; Indels 111; Gaps 20;

```
Qy    34  QAATDEEAVGFAHCDDLDFEHRQGLQWDILKAEGEPCCDEGAE--APVEGDIHYQRGSG   90

Db     2  EAPAPAEAAG---CEELLMDVMRPL---INEQNFDGSSDEEQTLLPMQK--HYQL---   50

Qy    91  APLPPSGSKDQVGGGGEFGGHDKPKITAWEAGWNVTNAIQGMFVLGLPYAILHGG-YLGL  149

Db    51  ---------DGQHG------ISFVQTLMHLLKGNIGTCLLGLPLAIKNAGIVLGP     90

Qy   150  FLIIFAAVVCCYTGKILIACLYEENEDGEVVRRDSYVAIANA--------CCAPRFP    199

Db    91  ISLVFIGIISVHCMHILVRCSHFLCQ-----RFKKSTLGYSDTVSFAMEASPWSCLQRQA 145

Qy   200  TLGGRVVNVAQIELVMTCILYVVVSGNLMYNSFPGL-------------PVS        239

Db   146  AWGRSVVDFFLVITQLGFCSVYIVFLAENVKQVHEGLLETTVVSNSSDLSQVCERRSVD  205

Qy   240  QKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARD--------  291

Db   206  LRVMLCFLPLLILIVFIRELKSLFVLSLANISMAASLVIIIYQYVVRSMPDPHNLPIVA  265

Qy   292  -WAWEKVKFYIDVKKFPISIGIIVFSYTS-QIFLPSLEGNMQQPSEFHCMMWTHIAACV  349

Db   266  GW----------KKYPLFFGTAVFAFRGIGVVLP-LENQMRESKRFPQALNIGMAIVTV  313
```

RESULT 10

```
Db   350  LKGLFLAVAYLTWADETKEVITDNLPGS--IRAVVNIFLVAKALLSYPLPFFAAVEVLEK  407

Qy   314  LYISLATLGVMCFRDEIKGSITLNLPQDMWLYQSVKILYSFGIFVTYSIQFYVPAEIIIP  373

Qy   408  SLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAG  467

Db   374  AVTARLHAKWXCICDFG-------IRSLLVSITCAGVLIPRLDIVISFVGAVSSST       423

Qy   468  LCFLLPSLFHIRLLWRK---LLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNA  523

Db   424  LALILPPLVEI-LTFSKDHYNVW-MVLKNISIAFTGFVGFLLGTYVTVEEII--YPTTA  478
```

RESULT 9
US-09-640-419C-25
; Sequence 25, Application US/09640419C
; Patent No. 6630615
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L
; APPLICANT: Crasta, Oswald R
; APPLICANT: Hu, Xu
; APPLICANT: Lu, Guihua
; TITLE OF INVENTION: DEFENSE-RELATED SIGNALING GENES AND METHODS OF USE
; FILE REFERENCE: 35718/199009 (5718-92)
; CURRENT APPLICATION NUMBER: US/09/640,419C
; CURRENT FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,656
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/206,405
; PRIOR FILING DATE: 2000-05-23
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-640-419C-25

Query Match   7.0%;   Score 192.5; DB 4; Length 418;
Best Local Similarity   23.3%; Pred. No. 2.5e-11;
Matches 95; Conservative 69; Mismatches 169; Indels 75; Gaps 16;

```
Qy   114  PKITAWEAGW------NVTNAIQGMFVLGLPYAILHGGY-LGLFLIIFAAVVCCYTGKIL  166

Db    19  PITASRNANWYSAFHNVT-AIVGAVGLGLPYAMSELGWGPGVVVLLSWVITLYTFWQM    77

Qy   167  IACLYEENEDGEVVRRDSYVAIANACCAPRFPTLGGRVVNVAQIELVMTCILYVVVSG   226

Db    78  I----EMHEMFEGKRF-DRYHELGQAAFGKK-- IGLYIVVPLQQLLVETSACIVMVTGG 129

Qy   227  ------NLMYNSFPGLPVSQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINI  279

Db   130  ESLKKIHQLSVGDYECRKLKVRHFIIIFASSQFVLSLLKNFNSIGGVSLVA--AVMSMSY 187

Qy   280  LVIAYCLSRAR------DWAWEKVKFYIDVKKFPISIGIIVFSYTS------QIFLPSL  326

Db   188  STIAWVASLTKGVANNVEYGYKRRNNTSVPLAFLGALGEMAFAYAGHNVVLEIQATIPST 247

Qy   327  -EGNMQQPSEFHCMMWTHIAACVLKGLFALVAYLTWADETKEVITDNL--PGSIRAVVN  383

Db   248  PENPSKRPMWKGAIVAYIIVAFCYFP--VALVGFWTFGNNVEENILKTLRGPKGLIIVAN 305

Qy   384  IFLVAKALLS---YPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSW--------G  432

Db   306  IFVIIHLMGSYQVYAMPVFDMIESV------------------MIKKWHFSPTRVLR    344

Qy   433  LTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL           480

Db   345  FTIRWTFVAATMGIAVALPHSALLSFFGGFIFAPTTYFIPCIIWLIL            392
```

US-09-178-093B-28
; Sequence 28, Application US/09178093B
; Patent No. 6660846
; GENERAL INFORMATION:
;   APPLICANT: Robert H. Edwards
;   APPLICANT: Richard J. Reimer
;   APPLICANT: Steve L. McIntire
;   APPLICANT: Erik M. Jorgenson
;   APPLICANT: Kim Schuske
;   TITLE OF INVENTION: Vesicular Amino Acid Transported
;   TITLE OF INVENTION: Composition and Method
;   FILE REFERENCE: 2002-0005.30
;   CURRENT APPLICATION NUMBER: US/09/178,093B
;   CURRENT FILING DATE: 2001-08-20
;   PRIOR APPLICATION NUMBER: 60/063,012
;   PRIOR FILING DATE: 1997-10-23
;   NUMBER OF SEQ ID NOS: 50
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 28
;   LENGTH: 504
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide (RVT3 polypeptide sequence)
US-09-178-093B-28

Query Match   7.0%;   Score 192;   DB 4;   Length 504;
Best Local Similarity   20.3%;   Pred. No. 3.7e-11;
Matches 102; Conservative   96; Mismatches 166;   Indels 138; Gaps 24;

Qy   57 GLQMDILKAEGEPCGD---EGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHD 112
                 ::: |||::|: :::| |||: :||
Db   16 GKHLEGLLPVGMPTADTQRAEDAQ-------HCGEGKGF-LQQSSSKE------------ 55

Qy   113 KPKITAWE-------AGWNVTNAIQGMFVLGLPYAILHGG-YLGLFLIIFAAVVCCYTGK 164
                 :|:|| |:||:||| |||||||||||||||| |:|||||||||||:|:|
Db   56 -PHFTDFEGKTSFGMSVFNLSNAIMGSGILGLAYAMANTGIIIFLFLLTAVALLSSYSIH 114

Qy   165 ILIIACLYEENEDGEVVRVRDSYVAIANACCAPRFPTLGGRVNVAQIIELV--MTCILYV 222
                 :|::   : :|:||:| :|: |||||||||||||||||||:|:   :||:||
Db   115 LLL-------KSSGIVGIR-AYEQLGYRA----FGTPGKLAAALAITLQNIGAMSSYIYI 162

Qy   223 VVSG-NLMYNSFPGLPVSQKSW-------SIIATAVLLPCAFLKNLKAV---------- 263
                 |:|: :| | :|:|| ::| ::|:||||:||:|||:||:|
Db   163 IKSELPLVIQTFLNLEKPTPVWIMDGNVLVILVSVIIILPLALMRQLGYLGYSSGFSLSC 222

Qy   264 --------SKFSLLCTLAHFVIN-------ILVI------------AYCLSRARDW 292
                         ||:||| ||||:|| :|:| ::|:
Db   223 MVFFLIAVIYKKFQVPCPLAHNLVNATGNFSHMVVVEEKSQLQSEPDTAEAFCTP----- 277

Qy   293 AWEKVKFPYIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNNWTHIAACVLKG 352
                  :: |||:||||||| |::|:||||:||||||||||| |:|||:|||
Db   278 -----SYFTLNSQTAYTIPIMAFAPVCHPEVLPIYTELKDPSK-RKMQHISNLSIAVMYV 331

Qy   353 LF--ALVAYLTWADETKEVITDNL----PGSIRAV-VNIFLVAKALLSYPLPFFAAVEV 404
                  |: |||||||||||:|:||| ||::|| ||:|:|||||:|||||||||||:|
Db   332 MYFLAALFGYLTFYDGVESELLHTYSKVDPFDVLLCVRVAVLIAVTLTVPIVLFPVRRA 391

Qy   405 LEKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVVFTL----LMAIYVPHFALLMGL 459
                 ::||| :: | :| |:||||:: |::||:||||||||
Db   392 IQQMLFQ--NQEF----------SW---LRHVIATGLLTCNLLVIFAPNILGIFGI 434

Qy   460 TGSLTGAGLCFLLPSLFHLRLL 481
                 ||||:|||||||||||||||:||
Db   435 IGATSAPCLIFIPPAIFYFRIM 456

RESULT 11
US-09-370-253-2
; Sequence 2, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
;   APPLICANT: Allen, Stephen M.

APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
;   LENGTH: 447
;   TYPE: PRT
;   ORGANISM: Hordeum vulgare
US-09-370-253-2

Query Match   6.7%;   Score 183.5;   DB 3;   Length 447;
Best Local Similarity 22.1%;   Pred. No. 2.4e-10;
Matches   96; Conservative   80; Mismatches 197;   Indels   61; Gaps 15;

Qy   119 WEAGWNVTNAIQGMFVLGLPYAILHGGY-LGLFLIIFAAVVCCYTGKILIACLYEENEDG 177
                 |:||:||||||||||||||||||||| ||||||||||||:| ||||||:|||
Db   39 WYSAFHNVTAMVGAGVLSLPYAMSELGWGPGIAVMTLSWIITVYT-----LWQMVEMH 91

Qy   178 EVVRVR--DSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPG 235
                  :|:|: |:|||||||:|:||||||||| ||:|:|
Db   92 EMVPGKRFDRYHELGQHAFGDK---LGLWIVVPQQLVVEVSLNIVYMVTGGNSL-KKFHD 147

Qy   236 LPVSQKSWSI-------IATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINLVIAYCLSR 288
                  |||:|:|:   :|||||||| |||::: ::||
Db   148 VICDGKCKDIKLTYFIMIFASVHFVLSQLPNFNSISGISLAAAVMSLSYSTIAWGASLHK 207

Qy   289 ARDWAWEKVKFYI-------DVKKFPISIGIIVFSYTS------QIFLPSLEGN-MQQP 333
                  ||: ::::||| |:|||||||||||:: |||||||:|
Db   208 GKE---ENVDYSLRASTTAGQVFGFLGGLGDVAFSYSGHNVVLEIQATIPSTPGNPSKKP 264

Qy   334 SEFHCMNWTHIAACVLKGLFALVAYLTWADETKE--VITDNLPGSIRAVVNIFLVAKAL 391
                 ||||||||||||||::: |:||||||:|| ::::
Db   265 MWKGVVVAYIIIAACYFP--VAFIGYWAFGNSVDDNILIIILNKPKWLIAMANMMVVHLI 322

Qy   392 LSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGLTL---RCALVVFTLLMA 447
                  ::::::: ||||||||:|| || |||::|||:||
Db   323 GSYQIYAMPVFDMMETFLVKK--LEFAP---------GTTLRLITRTIYVAFTMFIG 368

Qy   448 IYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLWRKLLMWHQVFFDVAIFVIGGICSVSG 507
                  ::|::||||||||:||||||||||||||||:||||||||||||::|||||||||||||||
Db   369 MSFPFFGGLIGFFGGLAFAPTTYFLPCIMWIIICKPRIFSLSWFTNWICIVLGVLLMVA 428

Qy   508 FVHSLEGLIEAYRT 521
                  :|::||||:|:||
Db   429 PIGGLRQIIISAKT 442

RESULT 12
US-09-919-497-67
; Sequence 67, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
;   APPLICANT: Mutter, George L.
;   TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
;   FILE REFERENCE: B0801/7225
;   CURRENT APPLICATION NUMBER: US/09/919,497
;   CURRENT FILING DATE: 2001-07-31
;   PRIOR APPLICATION NUMBER: US 60/221,735
;   PRIOR FILING DATE: 2000-07-31
;   NUMBER OF SEQ ID NOS: 100
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
;   LENGTH: 504
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-919-497-67

```
Query Match        6.6%;   Score 181.5;  DB 4;   Length 504;
Best Local Similarity  20.5%;   Pred. No. 4.7e-10;
Matches 101;  Conservative  97;  Mismatches 179;  Indels 115;  Gaps  23;

Qy   53  EHRQLQMDILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHD 112
             : | |||             :  :  :  :    ||:
Db   17  KHSEGLLPVITPMAGNQRVEDPARSCMEGKSFLQK--------SPSKE----------- 56

Qy  113  ---------AGWNVTNAIQGMFVLGLPYAILHGG-YLGLFLIIFAAVVCCYTGK 164
             ::|| | |  :  ||| :||||:||||:|  ||| |::|||||
Db   57  -PHFTDFEGKTSFGMSVVNLSNAIMGSGILGLAYAVANTGILLFLFLLTAVALLSSYSIH 115

Qy  165  ILIACLYEENEDGEVRVRDSYVAIANACCAPRFPTLGGRVVNVAQIELV-MTCILYV 222
             :|:||        |:::||||| |:  ::|  :::
Db  116  LIL--------KSSGVVGIR-AYEQLGYRA---FGTPGKLAAALAITLQNIGAMSSYLYI 163

Qy  223  VVSG-NLMVNSFPGLPVSQKSW-------SIIATAVLLPCAFLKNLKAV--SKFSLLC 270
             | ||  : ::  |: | |          :::::|| |||| : |:    ::|:||
Db  164  IKSELPLVIQTFLNLEBKTSDWYMNGNYLIVILVSVTIILPLALMRQLGYLGYSSGFSLSC 223

Qy  271  TLAHFVINLVIAY---C----------------LSRARDWAWEKVKFYIDVK-------- 304
             ||||||:|:|||    |                ::  | : : |:::
Db  224  -MVFFLIAVIYKKFHVPCPLPNFNNTTGNFSHV-EIVKEKVQLQVEPEASAFCTPSYFT 281

Qy  305  ---KFPISIGIIVFSYTSQIFLPSLEGNMQQPS--EFHCMNWTHIAACVLKGLFALVAY 359
             ||| :|||||   :  ::||  :  : ::          |:|||||:::::||||:||
Db  282  LNSQTAYTIPIMAFAFVCHPEVLPIYTELKDPSKKKMQHISNLSIAVMYIMYFLAALFGY 341

Qy  360  LTWADETKEVITDNL------PGSIRAV-VNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGS 414
             || ||| |: : ||       |||:|:| :||| || ::||||:::   ||::|||||
Db  342  LTFYNGVESELLHTYSKVDPFDVLILCVRVAVLTAVTLTVPIVLFPVRRAIQQMLFP--N 399

Qy  415  RAFFPACYSGDGRLKSWGLTLRCALVVFTL-----LMAIYVPHFALLMGLTGSLTGAGLC 469
             : :|||||||||:|||||:||:||||::||     ||||:|||||:|||||:|||
Db  400  QEF------------SW--LRHVLIAVGLLTCINLLVIFAPNILGIFGVIGATSAPFLI 444

Qy  470  FLIPSLFHLRLL 481
             |:| | |  |:||
Db  445  FIFPAIFYFRIM 456
```

RESULT 13
US-09-370-253-10
; Sequence 10, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
;   APPLICANT: Allen, Stephen M.
;   APPLICANT: Sakai, Hajime
;   APPLICANT: Thorpe, Catherine J.
;   TITLE OF INVENTION: Amino Acid Transporters
;   FILE REFERENCE: BB-1200
;   CURRENT APPLICATION NUMBER: US/09/370,253
;   CURRENT FILING DATE: 1999-08-09
;   EARLIER APPLICATION NUMBER: 60/097,222
;   EARLIER FILING DATE: August 20, 1998
;   NUMBER OF SEQ ID NOS: 18
;   SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-370-253-10

```
Query Match        6.4%;   Score 175;  DB 3;   Length 447;
Best Local Similarity  24.7%;   Pred. No. 1.9e-09;
Matches 112;  Conservative  66;  Mismatches 165;  Indels 110;  Gaps  26;

Qy  112  DKPKITAW------EAGW------NVTNAIQGMFVLGLPYAILHGGY-LGLFLIIFAAV 157
             : ||                     ||| ||| |||:||||||||| ||| |||||
Db   20  EQKKIDDWLPITSSRNAKWYSAFHVT-AMVGAGVLSLPYAMSELGWGPGIAVLVISWV 78

Qy  158  VCCYTGKILIACLYEENEDGEVRVR--DSYVAIANACCAPRFPTLGGRVVNVAQIELV 215
             ||  ||:|:|||||           ||||||||| |  |||:|||||:::
Db      (continues)
```

```
Db   79  ITLYT------LWQMVEMHEMVPGKRFPDRYHELQCHAFGDK--LGLWIVVPPQQLIVEV 128
             :||         :   | |:| :||||: ||||| ||| |||
Qy  216  MTCILYVVVSGNLMVNSFPGLPVSQKSWSI------IATAVLLPCAFLKNLKAVSKFSL 268
             || ||| | ||:: | |:||| |  |         ::|:|| ||||||||:: ::|:
Db  129  GVNIVMVTGGRSL-KKFHDVICDGKCKDIKLTFFIMIFASVHFVLSQLPNLNSISGVSL 187

Qy  269  LCTLAHFVINLVIAYCLSRARDWAW---EKVKFYIDV----KKFPI-SIG 311
             |||||| |: | |     ::       :| :||:|:         ||| |||
Db  188  AAA--------VMSLSYSTI--AWGASVDKGKM-VDVDYNLRATTTPGKVFGFFGALG 234

Qy  312  IIVFSYTS------QIFLPSL-EGNMQQPSEFHCMNWTHIAACVLKGLFALVALTWA 363
             ||:|          ||| |:||         :::         ::|||||||||||
Db  235  EVAFAYAGHNVVLEIQATIPSTPEKPSKKPMWKGVVAYIVVALCYFP--VALIGYWAFG 292

Qy  364  DETKE--VITDNLPGSIRAVVNIFLVAKALLS---YPLPFFAAVE-VLEKSLFQEGSRAF 417
             ||:        | || : : :|:|:|||||||    ||||| ::  ||:||||||:|||
Db  293  NSVDDNLITLNKPTWLISTANMMVVIHVIGSYQIYAMPVFDMIETVLVKKL-----K 345

Qy  418  FPACYSGDGRLKSWGLTL---RCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLP 473
             |||              |   :  ::|| :|||:|:|||||:|||||:|||||| ||
Db  346  FPP------------GJTLRLIARTIYVAFTMFVALTFPFPFGGLLGFFGGFAFAPTTYFLP 394

Qy  474  SLFHLRLLWRK--LLWHQVFPFDVAIFVIGGIC 503
             |:| | |     || :|| || |||::|| |:|:
Db  395  CIMWLAIYKPRFSLSW-----FTNMVCIVLGVC 423
```

RESULT 14
US-09-640-419C-5
; Sequence 5, Application US/09640419C
; Patent No. 6630615
; GENERAL INFORMATION:
;   APPLICANT: Bidney, Dennis L
;   APPLICANT: Crasta, Oswald R
;   APPLICANT: Hu, Xu
;   APPLICANT: Lu, Guihua
;   TITLE OF INVENTION: DEFENSE-RELATED SIGNALING GENES AND METHODS OF USE
;   FILE REFERENCE: 35718/199009 (5718-92)
;   CURRENT APPLICATION NUMBER: US/09/640,419C
;   CURRENT FILING DATE: 2000-08-17
;   PRIOR APPLICATION NUMBER: 60/149,656
;   PRIOR FILING DATE: 1999-08-18
;   PRIOR APPLICATION NUMBER: 60/206,405
;   PRIOR FILING DATE: 2000-05-23
;   NUMBER OF SEQ ID NOS: 28
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Helianthus annuus
US-09-640-419C-5

```
Query Match        6.3%;   Score 174;  DB 4;   Length 452;
Best Local Similarity  22.9%;   Pred. No. 2.4e-09;
Matches 100;  Conservative  70;  Mismatches 201;  Indels  66;  Gaps  17;

Qy  119  WEAGWNVTNAIQGMFVLGLPYAILHGGY-LGLFLIIFAAVVCCYTGKILIACLYEENEDG 177
             : | |||| ||| |||:|||||||||| ||| ||:|| :::
Db   43  WYAAFHNVTAMVGAGVLSLPYAMSELGWGPGVAVMVISWIVTVYT----LWQMVEMH 95

Qy  178  EVRVR--DSYVAIANACCAPRFPTLGGRVVNVAQIELVMTCILYVVVSGNLMYNSFPG 235
             ||          | : | || :||||:::      :|| || |  | | ::|||:||
Db   96  EMVPGKRFPDRYHELGQHAFGEK--LGLYIVVPPQQLVVEVSLCIIYMVTGGKSL-QKFHD 151

Qy  236  LPVSQKS------WSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLS 287
             | |:|         :||||| |||||||||::  :||: |||||||:|: : |
Db  152  LLVRQDEQKDIRLTFFIMIYGSVHFVLSHLPNFNSLSGISLAAAIMSLSYSTIAWGASLD 211

Qy  288  RA----RDWANEKVKFYIDVKKPPISIGIIVFSYTS------QIFLPSL-EGNMQQPSE 335
             |      | |  |||| | :             :         ||| |:|| :::
Db  212  KGVQPNVEYGYKAKSTGTVFNFLSAIGDMAFAYAGHNVVLEIQATIPSTPEKPSKGPMW 271

Qy  336  FHCMNWTHIAACVLKGLFALVALVAYLTWADETKE--VITDNLPGSIRAVVNIFLVAKALLS 393
             :::            ::|||||||||||  ||:      | || : : :|:|:|||||||
Db      (continues)
```

Fri Nov 12 12:23:21 2004

us-09-940-919-2.rai

```
Db   272  ::  :.  ||         |:||          ::||  ||    |:   ||:|:|
          KGVVVAYIVIALCYFP--VAFIGYWVFGNGVDDNILITLEKPTWLIAMANMFVVHVIGS 329

Qy   394  YPL---PFFAAVEVL--EKSLFQEGSRAFFPACYSGDGRLKSWGLITLRCALVFTLLMAI 448
                          |:  :||: |                              |  :|

Db   330  YQIYAMPVFDMMETLLVKKLNFAPSFMLRF---------------ITRNVVVAFTMFVGI 374

Qy   449  VYPHFALLMGLTGSLTGAGLCFLLPSLFHLRLL----WRKLLMHQVFFDVAIFVIGGICS 504
                     |:|   ||         |:|                |:|  ||:||:||:

Db   375  CPPFFGGLLGFFGGLAFAPTTYFLPCIMWLCICKPRRW-SLSW---FTNWICIVLGVVLM 430

Qy   505  VSGFVHSLEGLIEAYRT 521
                  :  |
Db   431  VVAPIGGLRSIIVQAKT 447
```

RESULT 15
US-09-640-419C-26
; Sequence 26, Application US/09640419C
; Patent No. 6630615
; GENERAL INFORMATION:
;   APPLICANT: Bidney, Dennis L
;   APPLICANT: Crasta, Oswald R
;   APPLICANT: Hu, Xu
;   APPLICANT: Lu, Guihua
;   TITLE OF INVENTION: DEFENSE-RELATED SIGNALING GENES AND METHODS OF USE
;   FILE REFERENCE: 3571B/199009 (5718-92)
;   CURRENT APPLICATION NUMBER: US/09/640,419C
;   CURRENT FILING DATE: 2000-08-17
;   PRIOR APPLICATION NUMBER: 60/149,656
;   PRIOR FILING DATE: 1999-08-18
;   PRIOR APPLICATION NUMBER: 60/206,405
;   PRIOR FILING DATE: 2000-05-23
;   NUMBER OF SEQ ID NOS: 28
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 26
;   LENGTH: 446
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-640-419C-26

Query Match   6.3%; Score 173; DB 4; Length 446;
Best Local Similarity 23.1%; Pred. No. 3e-09;
  Matches 99; Conservative 71; Mismatches 186; Indels 72; Gaps 19;

```
Qy   119  WEAGWNVTNAIQGMFVLGLPYAILHGGY-LGLFLIIFAAVVCCYTGKILIACLYEENEDG 177
               |: :|  |: :||||| :|: :|:|:|: |:|:           |:|:|
Db    39  WYSAFHNVTAMVGAGVLGLPYAMSQLGWGPGIAVLVLSWVITLYT------LMQMVEMH 91

Qy   178  EVVRVR--DSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPG 235
               |:     |:: |        :||   |:                |::|
Db    92  EMVPGKRFDRYHELGQHAFGEK---LGLYIVVPQOLIVEIGVCIIVYMVTGGKSL-KKFHE 147

Qy   236  L------PVSQKSWSIIATAVLLPCAFLKNLKAVS-KFSLLCTLAHFVIN----ILVIAY 284
          |:      |:  :       :         :|:                :|:
Db   148  LVCDDCKPIKLTYFIMIFASVHFVLSHLPNFNSISGSFSCCCRYVSQLLNNRMGIISKQR 207

Qy   285  CLSRARDWAWEKVKFVIDVKKFPISIGIIVPSYTS------QIFLPSL-EGNMQQPSBF 336
          |:                  :                   :||:|                |:|
Db   208  C-SRRRSIRLQSENNSRYVFNFFSGLGDVAFAYAGHNVVLEIQATIPSTPEKPSKGPMWR 266

Qy   337  HCMNWTHIAACVLKGLFALVAYLTWADETKEVITDNL--PGSIRAVVNIFLVAKALLS- 393
               :|    |||: |   ||||  ::|   |:||                |: |||  |
Db   267  GVIVAYIVIALCYFP--VALVGYYIFGNGVEDNILMSLKKPAWLIATANIFVVIHVIGSY 324

Qy   394  --YPLPFFAAVEVL--EKSLFQEGSRAFFPACYSGDGRLKSWGLITLRCALVFTLLMAIY 449
               :|:  ||::  |     |                               ||:|
Db   325  QIYAMPVFDMMETLLVKKLNFRPTTTLRF---------------FVRNFYVAATMFVGMT 369

Qy   450  VPHFALLMGLTGSLTGAGLCFLLPSLFHLRLL------WRKLLMHQVFFDVAIFV-- 498
                 |:|                  |:|                |:|
Db   370  FPFFGGLLAFFGGFAFAPTTYFLPCVIWLAIYKPKKYSLSWWAN--WVCIVFGLFMVLS 427
```

```
Qy   499  -IGGICSV 505
           |||: :|
Db   428  PIGGLRTI 435
```

Search completed: November 8, 2004, 18:56:47
Job time : 48 secs

This Page Blank (uspto)

```
                     GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     November  8, 2004, 18:44:35  ; Search time 173 Seconds
                                          (without alignments)
                                          1071.928 Million cell updates/sec

Title:           US-09-940-919-2
Perfect score:   2750
Sequence:        1 MATLLRSKLSNVATSVSNKS........SGFVHSLEGLIEAYRTNAED 525

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters:     1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
          1:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
          2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
          3:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
          4:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
          5:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
          6:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
          7:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
          8:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
          9:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
         10:  /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
         11:  /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
         12:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
         13:  /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
         14:  /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
         15:  /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
         16:  /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
         17:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
         18:  /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
         19:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
         20:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2750	100.0	525	9	US-09-871-503-2	Sequence 2, Appli
2	2750	100.0	525	9	US-09-940-919-2	Sequence 2, Appli
3	2750	100.0	525	11	US-09-972-211-34	Sequence 34, Appl
4	2750	100.0	525	11	US-09-972-211-105	Sequence 105, App
5	2750	100.0	525	14	US-10-239-566-2	Sequence 2, Appli
6	2750	100.0	525	15	US-10-096-625-34	Sequence 34, Appl
7	2750	100.0	525	15	US-10-096-625-105	Sequence 105, App
8	2750	100.0	525	16	US-10-789-241-10	Sequence 10, App
9	2746	99.9	525	11	US-09-972-211-30	Sequence 30, Appl
10	2746	99.9	525	15	US-10-096-625-30	Sequence 30, Appl
11	2746	99.9	525	15	US-10-096-625-38	Sequence 38, Appl
12	2746	99.9	525	15	US-10-096-625-38	Sequence 38, Appl
13	2742	99.7	525	11	US-09-972-211-36	Sequence 36, Appl
14	2742	99.7	525	15	US-10-096-625-36	Sequence 36, Appl
15	2739	99.6	525	11	US-09-972-211-32	Sequence 32, Appl
16	2739	99.6	525	11	US-09-972-211-40	Sequence 40, Appl
17	2739	99.6	525	15	US-10-096-625-32	Sequence 32, Appl
18	2739	99.6	525	15	US-10-096-625-40	Sequence 40, Appl
19	2731	99.3	525	11	US-09-972-211-103	Sequence 103, App
20	2731	99.3	525	15	US-10-096-625-103	Sequence 103, App
21	2700	98.2	525	11	US-09-972-211-104	Sequence 104, App
22	2700	98.2	525	15	US-10-096-625-104	Sequence 104, App
23	2700	98.2	1050	14	US-10-225-810-36	Sequence 36, App
24	2641	96.0	521	11	US-09-972-211-106	Sequence 106, App
25	2641	96.0	521	15	US-10-096-625-106	Sequence 106, App
26	1075	39.1	549	11	US-09-972-211-107	Sequence 107, App
27	1075	39.1	549	15	US-10-096-625-107	Sequence 107, App
28	858	31.2	462	14	US-10-369-493-5867	Sequence 5867, Ap
29	858	31.2	486	14	US-10-225-810-37	Sequence 37, Appl
30	376	13.7	577	15	US-10-425-114-65034	Sequence 65034, A
31	372.5	13.5	496	17	US-10-425-115-301852	Sequence 301852,
32	368	13.4	578	17	US-10-425-114-65416	Sequence 65416, A
33	364	13.2	555	17	US-10-425-115-357602	Sequence 357602,
34	357.5	13.0	531	17	US-10-739-930-9570	Sequence 9570, Ap
35	355	12.9	496	16	US-10-437-963-201705	Sequence 201705,
36	352.5	12.8	550	15	US-10-424-599-182680	Sequence 182680,
37	352	12.8	522	17	US-10-425-115-216062	Sequence 216062,
38	348.5	12.7	548	16	US-10-437-963-156455	Sequence 156455,
39	343	12.5	430	15	US-10-425-114-55549	Sequence 56549, A
40	335	12.2	706	16	US-10-437-963-145638	Sequence 145638,
41	324	11.8	370	15	US-10-080-334-136	Sequence 136, App
42	324	11.8	370	15	US-10-080-334-142	Sequence 142, App
43	323.5	11.8	438	17	US-10-425-115-297771	Sequence 297771,
44	323.5	11.8	475	17	US-10-425-115-297768	Sequence 297768,
45	323.5	11.8	487	15	US-10-425-114-68137	Sequence 68137, A

ALIGNMENTS

```
RESULT 1
US-09-871-503-2

; Sequence 2, Application US/0971503
; Patent No. US2002007675A1
; GENERAL INFORMATION:
; APPLICANT: George Christian Terstappen
; APPLICANT: Cinzia Felicita Sala
; TITLE OF INVENTION: POLYPEPTIDE
; FILE REFERENCE: QG 1013
; CURRENT APPLICATION NUMBER: US/09/871,503
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: GB 0013239.9
; PRIOR APPLICATION FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-871-503-2

Query Match             100.0%;  Score 2750;  DB 9;  Length 525;
Best Local Similarity   100.0%;  Pred. No. 2.1e-253;
Matches  525;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy         61 DILKABGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKABGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120

Qy        121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVCCYTGKILIACLYBENEDGEVV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVCCYTGKILIACLYBENEDGEVV 180
```

```
RESULT 2
US-09-940-919-2
; Sequence 2, Application US/09940919
; Patent No. US2002008239OA1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: No. US2002008239OA1el Human GABA Transporter Protein and Polynucl
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0228-USA
; CURRENT APPLICATION NUMBER: US/09/940,919
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/230,178
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 525
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-940-919-2

Query Match             100.0%;   Score 2750;   DB 9;   Length 525;
Best Local Similarity   100.0%;   Pred. No. 2.1e-253;
Matches   525;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy    61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120

Qy   121 AGWNVTNAIQGMFVLGLPYAILHGGYLGGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
         ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGWNVTNAIQGMFVLGLPVAILHGGYLGGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy   181 RVRDSYVAIANACCAPRFPTLGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RVRDSYVAIANACCAPRFPTLGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240

Qy   241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy   301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMWTHIAACVLKGLFALVAYL 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMWTHIAACVLKGLFALVAYL 360

Qy   361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy   421 CYSGDGRLKSWGLTLRCALVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CYSGDGRLKSWGLTLRCALVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy   481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
         |||||||||||||||||||||||||||||||||||||||||||||
Db   481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525


RESULT 3
US-09-972-211-34
; Sequence 34, Application US/09972211
; Publication No. US2004004824541
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glenda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US2004004824541el Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/235,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
```

; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-211-34

Query Match   100.0%; Score 2750; DB 11; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.1e-253;
Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120

Qy 121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy 181 RVRDSYVAIANACCAPRPPTLGGRVVNNVAQIIELVMTCILYVVSGNLMYNSFPGLPVSQ 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 RVRDSYVAIANACCAPRPPTLGGRVVNNVAQIIELVMTCILYVVSGNLMYNSFPGLPVSQ 240

Qy 241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy 301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360

Qy 361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy 421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy 481 LWRKLLMWHQVFPDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
          |||||||||||||||||||||||||||||||||||||||||||||
Db 481 LWRKLLMWHQVFPDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525

RESULT 4
US-09-972-211-105
; Sequence 105, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R

; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glenda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding T
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-211-105

Query Match   100.0%; Score 2750; DB 11; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.1e-253;
.Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120

Qy 121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy 181 RVRDSYVAIANACCAPRPPTLGGRVVNNVAQIIELVMTCILYVVSGNLMYNSFPGLPVSQ 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 RVRDSYVAIANACCAPRPPTLGGRVVNNVAQIIELVMTCILYVVSGNLMYNSFPGLPVSQ 240

Qy 241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy 301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360

RESULT 5
US-10-239-566-2

; Sequence 2, Application US/10239566
; Publication No. US20030069398A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New human GABA recetor
; FILE REFERENCE: VGATFRWS
; CURRENT APPLICATION NUMBER: US/10/239,566
; CURRENT FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-566-2

Query Match   100.0%;   Score 2750; DB 14; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.1e-253;
Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120

Qy   121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy   181 RVRDSYVAIANACCAPRPPTLGGRVVNVAQIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RVRDSYVAIANACCAPRPPTLGGRVVNVAQIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240

Qy   241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVNILVIAYCLSRADWAWEKVKFY 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVNILVIAYCLSRADWAWEKVKFY 300

Qy   301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMWTHIAACVLKGLFALVAYL 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMWTHIAACVLKGLFALVAYL 360

Qy   361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy   421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy   481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
           |||||||||||||||||||||||||||||||||||||||||||||
Db   481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525

RESULT 6
US-10-096-625-34

; Sequence 34, Application US/10096625
; Publication No. US20040068095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William P
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; APPLICANT: Ji, Weizhen
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141 CIP
; CURRENT APPLICATION NUMBER: US/10/096,625
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/972,211
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 34
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-096-625-34

Query Match   100.0%;   Score 2750; DB 15; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.1e-253;
Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7

US-10-096-625-105

; Sequence 105, Application US/10096625

; Publication No. US20040068095A1

; GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine B
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Elleman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; APPLICANT: Ji, Weizhen

; TITLE OF INVENTION: No. US20040068095Ale1 Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141 CIP
; CURRENT APPLICATION NUMBER: US/10/096,625
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/972,211
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 105
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-096-625-105

Query Match   100.0%; Score 2750; DB 15; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.1e-253;
Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120

Qy  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240

Qy  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYL 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYL 360

Qy  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy  481 LWRKLLWHQVFPDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
          |||||||||||||||||||||||||||||||||||||||||||||
Db  481 LWRKLLWHQVFPDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
```

Db   481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525

RESULT 8
US-10-789-241-10
; Sequence 10, Application US/10789241
; Publication No. US20040180332A1
; GENERAL INFORMATION:
;   APPLICANT: Powell, Douglas M.
;   TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;   TITLE OF INVENTION: AIDS AND HIV-RELATED DISORDERS USING 9145, 1725, 311, 837,
;   TITLE OF INVENTION: 58305, 156, 14175, 50352, 32678, 5560, 7240, 8865, 12396,
;   TITLE OF INVENTION: 12397, 13644, 19938, 2077, 1735, 1786, 10220, 17822, 33945,
;   TITLE OF INVENTION: 43748, 47161, 81982 OR 46777
;   FILE REFERENCE: MPI03-041P1RNOMNIM
;   CURRENT APPLICATION NUMBER: US/10/789,241
;   CURRENT FILING DATE: 2004-02-27
;   PRIOR APPLICATION NUMBER: US 60/454,202
;   PRIOR FILING DATE: 2003-03-12
;   PRIOR APPLICATION NUMBER: US 60/456,326
;   PRIOR FILING DATE: 2003-03-20
;   PRIOR APPLICATION NUMBER: US 60/465,240
;   PRIOR FILING DATE: 2003-04-24
;   PRIOR APPLICATION NUMBER: US 60/475,233
;   PRIOR FILING DATE: 2003-06-02
;   PRIOR APPLICATION NUMBER: US 60/478,952
;   PRIOR FILING DATE: 2003-06-16
;   PRIOR APPLICATION NUMBER: US 60/487,836
;   PRIOR FILING DATE: 2003-07-16
;   PRIOR APPLICATION NUMBER: US 60/500,111
;   PRIOR FILING DATE: 2003-09-04
;   NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 10
;   LENGTH: 525
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-789-241-10

Query Match   100.0%; Score 2750; DB 16; Length 525;
Best Local Similarity   100.0%; Pred. No. 2.1e-253;
Matches 525; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120

Qy 121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIPAAVVCCYTGKILIACLYEENEDGEVV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIPAAVVCCYTGKILIACLYEENEDGEVV 180

Qy 181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSPGLPVSQ 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSPGLPVSQ 240

Qy 241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy 301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYL 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYL 360

Qy 361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy 421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy 481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
         |||||||||||||||||||||||||||||||||||||||||||||
Db 481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525

RESULT 9
US-09-972-211-30
; Sequence 30, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
;   APPLICANT: Shimkets, Richard A
;   APPLICANT: Taupier Jr, Raymond J
;   APPLICANT: Burgess, Catherine E
;   APPLICANT: Zerhusen, Bryan D
;   APPLICANT: Mezes, Peter S
;   APPLICANT: Rastelli, Luca
;   APPLICANT: Malyankar, Uriel M
;   APPLICANT: Grosse, William M
;   APPLICANT: Alsobrook II, John P
;   APPLICANT: Lepley, Denise M
;   APPLICANT: Spytek, Kimberly Ann
;   APPLICANT: Li, Li
;   APPLICANT: Edinger, Shlomit
;   APPLICANT: Gerlach, Valerie
;   APPLICANT: Ellerman, Karen
;   APPLICANT: MacDougall, John R
;   APPLICANT: Gunther, Erik
;   APPLICANT: Millet, Isabelle
;   APPLICANT: Stone, David J
;   APPLICANT: Smithson, Glennda
;   APPLICANT: Szekeres Jr, Edward S
;   TITLE OF INVENTION: No. US20040048245A1e1 Human Proteins, Polynucleotides Encoding Th
;   TITLE OF INVENTION: Methods Of Using The Same
;   FILE REFERENCE: 21402-141
;   CURRENT APPLICATION NUMBER: US/09/972,211
;   CURRENT FILING DATE: 2001-10-05
;   PRIOR APPLICATION NUMBER: 60/238,325
;   PRIOR FILING DATE: 2000-10-05
;   PRIOR APPLICATION NUMBER: 60/238,323
;   PRIOR FILING DATE: 2000-10-05
;   PRIOR APPLICATION NUMBER: 60/238,400
;   PRIOR FILING DATE: 2000-10-05
;   PRIOR APPLICATION NUMBER: 60/238,397
;   PRIOR FILING DATE: 2000-10-06
;   PRIOR APPLICATION NUMBER: 60/238,401
;   PRIOR FILING DATE: 2000-10-06
;   PRIOR APPLICATION NUMBER: 60/238,379
;   PRIOR FILING DATE: 2000-10-06
;   PRIOR APPLICATION NUMBER: 60/238,402
;   PRIOR FILING DATE: 2000-10-06
;   PRIOR APPLICATION NUMBER: 30/238,384
;   PRIOR FILING DATE: 2000-10-06
;   PRIOR APPLICATION NUMBER: 60/238,373
;   PRIOR FILING DATE: 2000-10-06
;   PRIOR APPLICATION NUMBER: 60/238,372
;   PRIOR FILING DATE: 2000-10-06
;   PRIOR APPLICATION NUMBER: 60/238,383
;   PRIOR FILING DATE: 2000-10-06
;   PRIOR APPLICATION NUMBER: 60/238,382
;   PRIOR FILING DATE: 2000-10-06
;   PRIOR APPLICATION NUMBER: 60/275,892
;   PRIOR FILING DATE: 2001-03-14
;   PRIOR APPLICATION NUMBER: 60/296,860
;   PRIOR FILING DATE: 2001-06-08
;   NUMBER OF SEQ ID NOS: 198
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
;   LENGTH: 525
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-211-30

Query Match   99.9%;   Score 2746; DB 11; Length 525;
Best Local Similarity 99.8%;   Pred. No. 5.1e-253;
Matches 524; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM  60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120

Qy  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240

Qy  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360

Qy  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy  481 LWRKLLWHQVFPFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
        |||||||||||||||||||||||||||||||||||||||||||||
Db  481 LWRKLLWHQVFPFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
```

RESULT 10
US-09-972-211-38

; Sequence 38, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glenda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 38
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-972-211-38

Query Match   99.98%;   Score 2746; DB 11; Length 525;
Best Local Similarity 99.8%;   Pred. No. 5.1e-253;
Matches 524; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

```
Qy    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM  60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120

Qy  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240

Qy  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360

Qy  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TWADETKEAITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy  481 LWRKLLWHQVFPFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
```

Query Match   99.9%; Score 2746; DB 15; Length 525;
Best Local Similarity   99.8%; Pred. No. 5.1e-253;
Matches 524; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

```
Qy    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGGGAPLPPSGSKDQVGGGGEFGGHDKPKITAWE 120

Qy  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKIIIACLYEENEDGEVV 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKIIIACLYEENEDGEVV 180

Qy  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240

Qy  241 KSWSIIATAVLLPCAFLNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
        |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 KSWSIIATAVLLPCAFLNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360

Qy  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy  481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
        ||||||||||||||||||||||||||||||||||||||||||||
Db  481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
```

RESULT 12
US-10-096-625-38
; Sequence 38, Application US/10096625
; Publication No. US20040068095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glenda
; APPLICANT: Szekeres Jr, Edward S
; APPLICANT: Ji, Weizhen
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141 CIP
; CURRENT APPLICATION NUMBER: US/10/096,625

```
Db   421 CYSGDGRLKSWGLTLRCALVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   481 LWRKLLWHQVFPDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
                |||||||||||||||||||||||||||||||||||||||||||||

Db   481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
```

RESULT 13
US-09-972-211-36
; Sequence 36, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1e1 Human Proteins, Polynucleotides Encoding T
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-211-36

; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/972,211
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-096-625-38

Query Match   99.9%; Score 2746; DB 15; Length 525;
Best Local Similarity 99.8%; Pred. No. 5,1e-253;
Matches 524; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAAATDEEAVGFAHCDDLDFEHRQGLQM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGBFGGHDKPKITAWE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGBFGGHDKPKITAWE 120

Qy   121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKIIIACLYBENEDGEVV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKIIIACLYBENEDGEVV 180

Qy   181 RVRDSYVAIANACCAPRFPTLGGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RVRDSYVAIANACCAPRFPTLGGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240

Qy   241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy   301 IDVKKPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMWNTHIAACVLKGLFALVAYL 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 IDVKKPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMWNTHIAACVLKGLFALVAYL 360

Qy   361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TWADETKEAITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy   421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Query Match        99.7%;    Score 2742;    DB 15;    Length 525;
Best Local Similarity    99.6%;    Pred. No. 1.2e-252;
Matches 523;    Conservative 0;    Mismatches 2;    Indels 0;    Gaps 0;

```
Qy    1  MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy   61  DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120

Qy  121  AGWNVTNAIQGMFVLGLPYAILHGGYIGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AGWNVTNAIQGMFVLGLPYAILHGGYIGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy  181  RVRDSYVAIANACCAPRFPTLGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
         |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RVRDSYAAIANACCAPRFPTLGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240

Qy  241  KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy  301  IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360

Qy  361  TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
         |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  TWADETKEAITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy  421  CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
```

RESULT 14
US-10-096-625-36

Query Match   99.7%;   Score 2742;   DB 11;   Length 525;
Best Local Similarity   99.6%;   Pred. No. 1.2e-252;
Matches 523;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

```
Qy    1  MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60

Qy   61  DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120

Qy  121  AGWNVTNAIQGMFVLGLPYAILHGGYIGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AGWNVTNAIQGMFVLGLPYAILHGGYIGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy  181  RVRDSYVAIANACCAPRFPTLGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
         |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RVRDSYAAIANACCAPRFPTLGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240

Qy  241  KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy  301  IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYL 360

Qy  361  TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
         |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  TWADETKEAITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

Qy  421  CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy  481  LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
         |||||||||||||||||||||||||||||||||||||||||||||
Db  481  LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
```

Query Match   99.6%; Score 2739; DB 11; Length 525;
Best Local Similarity 99.6%; Pred. No. 2.4e-252;
Matches 523; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM  60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE  120

Qy  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV  180

Qy  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ  240

Qy  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY  300

Qy  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYL  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYL  360

Qy  361 TWADETKEVITDNLPGSIRAVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TWADETKEAITDNLPGSIRAVNIFPVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA  420

Qy  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL  480

Qy  481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
          |||||||||||||||||||||||||||||||||||||||||||||
Db  481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
```

Search completed: November 8, 2004, 18:59:58
Job time : 183 secs

Db 421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480

Qy 481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
          |||||||||||||||||||||||||||||||||||||||||||||
Db 481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525

RESULT 15
US-09-972-211-32
; Sequence 32, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 32
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-211-32

This Page Blank (uspto)

```
              GenCore version 5.1.6
       Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:  November 8, 2004, 18:41:08 ; Search time 42 Seconds
                         (without alignments)
                         1202.710 Million cell updates/sec

Title:          US-09-940-919-2
Perfect score:  2750
Sequence:       1 MATLLRSKLSNVATSVSNKS.........SGFVHSLEGLIEAYRTNAED 525

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:        283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    PIR_79:*
          1:  Pir1:*
          2:  Pir2:*
          3:  Pir3:*
          4:  Pir4:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	858	31.2	462	2	S42372	hypothetical prote
2	858	31.2	486	2	T42254	amino acid permeas
3	320.5	11.7	516	2	T48238	hypothetical prote
4	318.5	11.6	529	2	E84813	hypothetical prote
5	317	11.5	543	2	T48239	hypothetical prote
6	306	11.1	449	2	H88022	protein T27A1.5 [i
7	298	10.8	607	2	T26845	hypothetical prote
8	295.5	10.7	436	2	T05653	amino acid transpo
9	285	10.4	571	2	T06737	hypothetical prote
10	272	9.9	426	2	T51506	hypothetical prote
11	252	9.9	481	2	JC7961	proton-coupled ami
12	241	8.8	602	2	S55188	hypothetical prote
13	236	8.6	423	2	T49959	hypothetical prote
14	223	8.1	484	2	T34016	hypothetical prote
15	222.5	8.1	472	2	JC7626	amino acid transpo
16	211	7.7	460	2	F85544	protein F59B2.2 [i
17	209	7.6	381	2	T02589	hypothetical prote
18	208	7.6	487	2	JC7328	amino acid transpo
19	202	7.3	490	2	S49792	probable membrane
20	201	7.3	713	2	S58251	probable membrane
21	198.5	7.2	489	2	A96837	hypothetical prote
22	196	7.1	389	2	S31123	hypothetical prote
23	186	6.8	476	2	C96505	probable amino aci
24	185.5	6.7	420	2	T39523	probable amino aci
25	184.5	6.7	467	2	T26705	probable transmemb
26	184.5	6.7	656	2	T38741	hypothetical prote
27	181.5	6.6	434	2	T15799	major facilitator
28	177	6.4	890	2	T21000	hypothetical prote
29	176	6.4	440	2	F86385	probable lysine an
30	174	6.3	505	2	B88206	protein F21D12.3 [
31	172.5	6.3	503	2	T26524	hypothetical prote
32	172	6.3	480	2	S50525	hypothetical prote
33	172	6.3	692	2	S37976	hypothetical prote
34	171.5	6.2	454	2	T02100	hypothetical prote
35	169	6.1	621	2	T20863	hypothetical prote
36	167.5	6.1	494	2	T16658	hypothetical prote
37	165	6.0	441	2	C86378	protein F21J9.6 [i
38	162.5	5.9	475	2	T00620	probable amino aci
39	162	5.9	481	2	T23131	hypothetical prote
40	160.5	5.8	448	2	S50622	hypothetical prote
41	153	5.6	462	2	T15052	amino acid permeas
42	151	5.5	476	2	H96802	probable amino aci
43	150	5.5	450	2	E96738	hypothetical prote
44	150	5.5	486	2	T10100	amino acid transpo
45	149	5.4	411	2	C57479	amino acid transpo

ALIGNMENTS

```
RESULT 1
S42372
hypothetical protein T20G5.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S42372
R;Smith, A.
submitted to the EMBL Data Library, March 1994
A;Reference number: S42368
A;Accession: S42372
A;Molecule type: DNA
A;Residues: 1-462 <SMI>
A;Cross-references: UNIPROT:P34579; EMBL:Z30423; NID:g458479; PID:g458484
C;Genetics:
A;Introns: 21/1; 91/3; 176/1; 227/2; 403/1; 429/2
C;Keywords: transmembrane protein

Query Match             31.2%;  Score 858;  DB 2;  Length 462;
Best Local Similarity   40.4%;  Pred. No. 1.1e-63;
Matches  180;  Conservative   96;  Mismatches  141;  Indels  28;  Gaps   8;

Qy   75 AEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWEEAGWNVTNAIQGMFV 134
            |: |:::|                      |||         :|||||:|||||||:
Db   39 SEQPQKDDINKQ--------EEAKDD--GHGE----ASEPISALQAAWVTNAIQGMFI 83

Qy  135 LGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVVRVRDSYVAIANACC 194
            |||  ||| ::  :||||||| ||  :|: ||:|||:  ||  ||||:
Db   84 VGLPIAVKVGGWWSIGAMVGVAYVCYWTGVLLIECLYENG----VKKRKTYREIAD-FY 137

Qy  195 APRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQKSWSIIATAVLLPC 254
            : ||| :||  :|:|  |::|: :|:|||
Db  138 KPGF---GKWVLAAQLTELLSTCIIYLVLAADLLQSCFPS--VDKAGWMMITSASLLTC 191

Qy  255 AFLKNLKAVSGKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKPYIDVKKFPISIGIIV 314
            |||    :  ||| |||:     ::|:::  |||
Db  192 SFLDDLQIVSRLSFFNAISHIVNLIMVLYCLSPVSQWSFSTITFSLNINTLPTIVGMVV 251

Qy  315 FSYTSGIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYLTWADETKEVITDNL 374
            |:| :|||  ||||:  ||||||||  |  ||  :|  ||| :|:: ||
Db  252 FGYTSHIFLPNLEGNMKNPAQFNVMLKWSHIAAAVFKVVFGMLGFLTFGELTQEISNSL 311

Qy  375 PG-SIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGL 433
            || : :  ||||||  ||||||| ||||: ||: :: :: :||
Db  312 PNQSFKILVNLILVVKALLSYPLPFYAAVQLLKNNLFLGYPQTPFTSCYSPDKSLREWAV 371

Qy  434 TLRCALVVFTLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLWRKLLWHQVFFD 493
            ||| :|: : :|::   ||| ||:: |  || ::  |:|||  ||  |:: |:|
Db  372 TLRIIIVLFTLFVALSVPYLVELMGLVGNITGTMLSFIWPALFHLYIKEKTLNNFEKRFD 431

Qy  494 VAIFVIGGICSVSGFVHSLEGLIEA 518
            ::|| ||| ::|| ||: ::
Db  432 QGIIIMGCSVCISGVYFSSMELLRA 456
```

RESULT 2
T42254
amino acid permease homolog.- Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T42254
R;McIntire, S.L.; Reimer, R.J.; Schuske, K.; Edwards, R.H.; Jorgensen, E.M.
Nature 389, 870-876, 1997
A;Title: Identification and characterization of the vesicular GABA transporter.
A;Reference number: Z22129; MUID:98007977; PMID:9349821
A;Accession: T42254
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-486 <MCI>
A;Cross-references: UNIPROT:P34579; EMBL:AF031935; NID:g2642593; PIDN:AAB87066.1; PID:g2
C;Genetics:
A;Note: unc-47

Query Match   31.2%;   Score 858; DB 2; Length 486;
Best Local Similarity   40.4%;   Pred. No. 1.1e-63;
Matches 180;   Conservative   96;   Mismatches 141;   Indels   28;   Gaps   8;

Qy   75 AEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWEAGWNVTNAIQGMFV 134
               :|:|| ||||||||||||||||||:::: :| | |||||||||||||||||||
Db   63 SEQPQKDDINKQ---------------EEAKDD--GHGE----ASEPISALQAAWNVTNAIQGMFI 107

Qy   135 LGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVVRVRDSYVAIANACC 194
               |||||||||||||||||||||||||||||||||||||||| ::|| :| ||
Db   108 VGLPIAVKVGGWNSIGAMVGVAYVCYWTGVLLIECLYENG-----VKKRKTYREIAD-FY 161

Qy   195 APRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQKSWSIIATAVLLPC 254
               |:|||||| || |:|||||:|| |:| | | || ||| ::: | ::::|||||||
Db   162 KPGF----GKWVLAAQLTELLSTCIIYVLVLAADLLQSCFPS--VDKAGWMMITSASLLTC 215

Qy   255 AFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFYIDVKKFPISIGIIV 314
               |||||||||||||||||||||||||||:|||| : || |||::| | |||||||||
Db   216 SFLDDLQIVSRLSFFNAISHLIVNLIMVLYCLSFVSQWSFSTITTSLNINTLPTIVGMVV 275

Qy   315 FSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAYLTWADETKEVITDNL 374
               |||| ||||||||||| : | || |||:|||||||||||||||||||||||||||||
Db   276 FGYTSHIFLPNLEGNMKGNPAQFNVMLKWSHIAAAVFKVVFGMLGFITFGELTQEEISNSL 335

Qy   375 PG-SIRAVVNIFIVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGL 433
               || || :|||:||||||||| |||||||||||||||:|||||||||||||||||
Db   336 PNQSFKIIVNLILIVVKALLSYPIPFYAVAVQLLKNNLFLGYPCQTPFTSCYSPDKSLREWAV 395

Qy   434 TLRCALVVFTLLMAIYPHFALLMGLTGSLTGAGLCFLLPSLFHLRILWRKLLWHQVFFD 493
               ||:|||||||| ||||||||||||| | | ||||||| || || ||||| || ||
Db   396 TLRIILVLFTLFVALSVPYLVELMGLVGNITGTMLSFIWPALFHLYIKEKTLNNFEKRFD 455

Qy   494 VAIFVIGGICSVSGFVHSLEGLIEA 518
               | :|:| ||:||||:|||||:
Db   456 QGIILMGCSVCISGVYFSSMELLRA 480

RESULT 3
T48238
hypothetical protein T7H20.220 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48238
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24488
A;Accession: T48238
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <BEV>
A;Cross-references: UNIPROT:Q9LZL5; EMBL:AL162508
A;Experimental source: cultivar Columbia; BAC clone T7H20
C;Genetics:

A;Map position: 5
A;Introns: 53/2; 144/1; 215/3; 276/1; 347/2; 392/3; 450/1; 464/3; 485/3
A;Note: T7H20.220
C;Superfamily: Arabidopsis amino acid transport protein I

Query Match   11.7%;   Score 320.5; DB 2; Length 516;
Best Local Similarity   24.1%;   Pred. No. 8.2e-19;
Matches 101;   Conservative   77;   Mismatches 166;   Indels   75;   Gaps   12;

Qy   124 NVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVVRVR 183
               |:||||||||:||| :||: || |||| ::| | ||||||||||| | | || |
Db   143 NGINVLCGVALLTMPYAVKEGGWLGLFLFSFGIIITYTGILLKRCL--ENSPG------ 194

Qy   184 DSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPV---- 238
               :|||| || || :| || | | : ::| :||::: | || ||||
Db   195 -------IHTYPDIGQAAPGTTGRILVSASCVEYIIMMSDNLSRMFPNTSLYINGF 243

Qy   239 --SQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAW- 294
               : |||| :||| :| :| :| ||||| : : | ||: | ||
Db   244 SLDSTQVFAITTLIVLPTVWLKDLSLLSYLS---------GVISSILLALCLF---WAGS 292

Qy   295 -EKVKFYI----DVKKFPISIGIVFSYTSQIFLPSLEGNMQQPSEF--------HCM 339
               ||:|: |||||||||||: ||| |:| || ||||| : |
Db   293 VDGVGFHISGQALDITNIPVAIGIYGFGFGSHSVFPNIYSSMKEPSKFPTVLLISFAFCT 352

Qy   340 MNWTHIAACVLKGLFALVAYLTWADETKEVITDNLPGSIRA--VVNIPLVAKALLSYPLP 397
               | | | : ||| | ::|: | | || :| |||| :|| | || |||:||
Db   353 LFYIAVAVC----------GFTMFGDAIQSQFTLNMPPHFTSSKIAVWTAVVTPMTKYALT 403

Qy   398 FFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVVFTLLMAIYPHFALLM 457
               | || ||| | | : :| | | :: |:| | ::|:| |||||||
Db   404 ITPVMLSLEB-LIPSSSRKM------------RSKGVSMLFRTILVLSTLVVALTVPFFATVA 453

Qy   458 GLTGSLTGAGLCFLLPSLFHLRLLWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLI 516
               ||| ||||| ||||:| ||:| : : | | | :| |: ||: || |:| |:
Db   454 ALIGSFIAMLIALIFPCLCYISIMKGRLTNFQIGICILIVIGISGGCGTYSAIARLI 512

RESULT 4
E84813
hypothetical protein At2g39130 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: E84813
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84813
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-529 <STO>
A;Cross-references: UNIPROT:Q9ZU25; GB:AE002093; NID:g6598488; PIDN:AAC79623.2; GSPDB:GN
C;Genetics:
A;Gene: At2g39130
A;Map position: 2

Query Match   11.6%;   Score 318.5; DB 2; Length 529;
Best Local Similarity   22.0%;   Pred. No. 1.2e-18;
Matches 120;   Conservative   96;   Mismatches 189;   Indels 141;   Gaps   19;

Qy   16 VSNKSQAKMSGMFARMG-----------------------------FQAATDEEAVGFAHCDDLDFEHR--- 55
               : ||: |: || : |
Db   77 LGNNSWTRFGSSFLSSGLIRRHTPBSLPTVTKPLLEEQADEQALP--------KHRLSS 127

Qy   56 QGLQMDILKAEGEPCDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPK 115
               : : | || || | |:
Db   128 QGLLSPIPSRRGSMRKDE-KSSMVSHEIPMSRNS-----------------SYG----- 163

Qy   116 ITAWEAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIACLYEENE 175
               ||: || : :| | |:| :|||||::: | |: |::|

Db   379 CTVLYIAVAVCGYTMFGEAVESQFTLNMPKHFFPSKVAVWTAVITPMTKYALTITPIVMS 438
             :: | | :: |:|:|
Qy   164 ---QAVLNGLNVLCGVGILSTPYAAKEGGWLGLMILFVYGLLSFYTGILLRYCLDSESD 219

Db   405 LEKSLFQEGSRAFFPACYSGDGRLKSWGLTL--RCALVFTLLMAIVPHFALLMGLTGS 462
             :| |:|:| ||
Qy   176 DGEVVRVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPG 235

Db   439 LEELI------------PTAKMRSRGVSILFRTMLVTSTLVVALSVPFFAIVAALIGS 484
                :::
Qy   220 ---LETYPDIGQAA---FGTTGRIFVSIVLYLELYACCVEYIILESDNLSSLYPN 268

Qy   463 LTGAGLCFLLPSLFHLRLLWRKLLWHQV---FPDVAIFVIGGICSVSGFVHSL 512
             :| |:|:|
Qy   236 LPVS------QKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLS 287

Db   485 FLAMLVALIFPCLCYLSILKGKLSNTQIGLCIFIIVFGVVSGCCGTYSAISRL 537
             :| |:|:|
Qy   269 AALSIGGFQLDARHLFALLTTLAVLPTVWLRDLSVLSYSIS-----AGGVIASVLVVLCL- 322

RESULT 6
H88022
protein T27A1.5 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: H88022
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A;Reference number: A75000; MUID:99069913; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_el
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A;Accession: H88022
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-449 <STO>
A;Cross-references: UNIPROT:O17275; GB:chr_II; PIDN:AB71045.1; PID:g2429516; GSPDB:GN00
C;Genetics:
A;Gene: T27A1.5
A;Map position: 2
C;Superfamily: Arabidopsis amino acid transport protein I

Query Match   11.1%;   Score 306;   DB 2;   Length 449;
Best Local Similarity 23.4%;   Pred. No. 1.1e-17;
Matches   107;   Conservative   91;   Mismatches   219;   Indels   40;   Gaps   15;

Qy   85 YQRGSGAPLPPSGSKDQVGGGGEFGG--HDKPKITAWEAGNVTNAIQGMFVLGLPYAIL 142
             :| |:|:|
Db   6 YDTARGNELPLRGSHQPTMGEMFASRVRDSRSITADQALIHMIKVMGTMGLSLPLAFK 65

Db   143 HGG-YLGLFLIIFAAVCCYTGKILI-----ACLYEENEDGEVVRVRDSVVAIANACCAP 196
             :| |:|:|
Db   66 HSGIWLGLILLCFICLICIYCTRQLIFGQHYITFIKREQRMDYANVMRSAVELGPAWIRG 125

Qy   197 RFPTLGGRVNVAQII-ELVMTCILYVVVSGNL--MYNSFPGLPVSQKSWSIIATAVLLP 253
             :| |:|:| |||
Db   126 H-GYLFKQMVNINMFVAQFGPCCVFVFMADNLKQFFDQTSSIHISQAGNIAL---LLIP 181

Qy   254 CAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDW-AWEKVKFYIDVKKFPISIGI 312
             :| |:|:|
Db   182 ISALCTIRELKALAPLAAVANFVYIIAVVVLADLFSDWQPLDSLPAFGAVENLPLFFGT 241

Db   313 IVFSYTSQIFLPSLEGNMQQPSEF--HCMNWTHIAACVLKGLFALVAVLTWADETKEV 369
             :| |:|:|
Db   242 VMFAFEGVAVVLPIENQMNEPIHFITPNGVLNTSCILVLLVYMTVGFFGFLRYGNDIKDT 301

Qy   370 ITDNLPGS-IRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRA-FFPACYSGDGR 427
             :| |:|:|
Db   302 LTLNLPQTPYQAIKVMFVLCILVSYPLQFYVPNERVEKVKRKVVEAKQEPMIYA---- 357

Qy   428 LKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLW---- 482
             :| |:|:|
Db   358 IRFGGVLLTCA-------MAQLIPHLALFISLVGSVAGTSLTLVPPPLIELLCSYSKQEL 410

Qy   483 RKLLWHQVFPDVAIFVIGGICSVSGFVHSLEGLIEAY 519
             :| |:|:|
Db   411 TKWVWIR---NIGLMAFAMVGFTTGTYASMVQIIEAF 444

RESULT 7
T26845

hypothetical protein Y43F4B.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26845
R;Matthews, L.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z20276
A;Accession: T26845
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-607 <WIL>
A;Cross-references: UNIPROT:O45936; EMBL:AL021481; PIDN:CAA16336.1; CESP:Y43F4B.7
A;Experimental source: clone Y43F4B
C;Genetics:
C;Gene: CESP:Y43F4B.7
A;Introns: 53/3; 80/3; 110/2; 141/2; 169/3; 229/1; 261/3; 300/2; 320/2; 355/3; 377/1; 41

Query Match   10.8%;   Score 298; DB 2;   Length 607;
Best Local Similarity 22.8%;   Pred. No. 7.5e-17;
Matches   100;   Conservative   94;   Mismatches 186;   Indels   58;   Gaps   17;

Qy   72 DEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWEAGWNVTNAIQG 131
               || ||| |:|:
Db   13 DNEPEAPGDTEVQTPRRNTVSEDTSLFQDRL----------PTENSLTPEQAFIHMVKAMLG 64

Qy   132 MFVLGLPYAILHGG-YLGLFLIIFAAVVCCY-TGKILIACLYEENEDGEVVRVRD--SYV 187
               ||| || : || |||
Db   65 TGLLSLPLAFKHSGLFLGLILTVLICLICLYCMRQVVFAAHFVCNRNG-----RDLIDYA 119

Qy   188 AIANAC--CAPRFPTLGG---RVVNVAQII-ELVMTCILYVVVSGNL-MYNSPPGLPV 238
              : |||| :|:   :|: ||
Db   120 NIMRGAVEMGPPWIKRNGYFFKQLVNVNMFISQLGFCCVIFVFMADNLEDFFNNTSIHL 179

Qy   239 SQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVK 298
               ||| || :
Db   180 SKAVWMLLLIPMLSICSIRRLSILAPFAMANVVVVAVAVLFFFLSDLRPIS--SLP 237

Qy   299 FYIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHI--AAC-VLKGLFA 355
               :| | | ||
Db   238 WFGKATDLPLFFGTVMFAFEGVAVIMPIENRMQSP--HAFISWNGVLNSSCLVVLAIFS 294

Qy   356 LV---AYLTWADETKEVITDNLPGS-IRAVVNIFLVAKALLSYPLPFFAAVEVLEKSL-- 409
                     ||| ||| :|: ||:|| ||
Db   295 VTGFYGYLSLGNDVKDTATLNLPMTPYQTIKLMFVACIMISYPLQFYVMERIEKWITR 354

Qy   410 ---FQEGSRAFFPACYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLITGA 466
                                                   || |: |||
Db   355 KIPVDKQTLYIYIARYSG---------------VILTCAIAELIPHLALFISLIGAFSGA 399

Qy   467 GLCFLLPSLFHLRLLWRK 484
                 :|| :||
Db   400 SMALLFPPCIELLTSYAK 417

RESULT 8
T05653
amino acid transport protein homolog F22I13.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05653
R;Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15420
A;Accession: T05653
A;Molecule type: DNA
A;Residues: 1-436 <BEV>
A;Cross-references: UNIPROT:Q9SVG0; EMBL:AL035539
A;Experimental source: cultivar Columbia; BAC clone F22I13
C;Genetics:
A;Map position: 4
A;Note: F22I13.20
C;Superfamily: Arabidopsis amino acid transport protein I

Query Match   10.7%;   Score 295.5; DB 2;   Length 436;
Best Local Similarity 25.8%;   Pred. No. 8.2e-17;
Matches   116;   Conservative   79;   Mismatches 188;   Indels   67;   Gaps   14;

Qy   93 LPPSGSKDQVGGGEFGGHDKPKITAWEAGWNVTNAIQGMFVLGLPYAILHGGYL-GLFL 151
               |:| |||| |||| |:|:
Db   16 IPPPARED-----SPLLGKGPFLSSQFKTFANVFIAVVGAGVLGLPYAFKRTGWLMGVLL 70

Qy   152 IIFAAVVCCYTGKILIACLYEENEDGEVVRVRDSYVAIANACCAPRFPTLGGRVVNVAQI 211
               :: |: | || : |:||
Db   71 LVSVSVLTHHCMLLVYTRRKLDSFNAGISKIGSFGDLGFAVCG---SLGRIVDLFII 126

Qy   212 IELVMTCILYVVVSGNLMYN-SFPGLP-----------------VSQKS---WSIIATAV- 250
               || :|:   :
Db   127 LSQAGFCVGYLIFIGTTANLSDPESPTSLRHQFTRLGSEFLGVSSKSLYINGCFPPQLG 186

Qy   251 LLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFYIDVKKFPISI 310
               | || |:
Db   187 LNSIKTLTHLAPLSIFADIVDLGAMAVVIVEDSMIILKQR----PDVVAFGGMSLFLYGM 242

Qy   311 GIIVFSYTSQIFLPSLEGNMQQPSEFHCMM------NWTHIAACVLKGLFALVAVLTWA 363
               |:|:| :   : |:
Db   243 GVAYYSFEGVGMVLPLESEMKDKDFGKVLALGMGFISLIYIA------FGILGYLAFG 295

Qy   364 DETKEVITDNL-PGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACY 422
               | ||: | : ||| |
Db   296 EDTMDIITANLGAGLVSTVVQLGICINLFFTFPLMMNPVFEIVB--------RRFSRGMY 347

Qy   423 SGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLITGAGLCFLLPSLFHLRLLW 482
               | || ||: ||||
Db   348 S------AW---LRWVLVLAVTLVALFVPNFADFLSLVGSSTCCVLGFVLPALFHLLVFK 398

Qy   483 RKLLWHQVFFDVAIFVICGGICSVSGFVHSL 512
                 :| : ||: |||
Db   399 EEMGWLQWSDTAIVLGVVLAVSGTWSSL 428

RESULT 9
T06737
hypothetical protein F28P10.190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06737
R;Quetier, F.; Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigu
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15793
A;Accession: T06737
A;Molecule type: DNA
A;Residues: 1-571 <QUE>
A;Cross-references: UNIPROT:Q9SV32; EMBL:AL049655; GSPDB:GN00061; ATSP:F28P10.190
A;Experimental source: cultivar Columbia; BAC clone F28P10
C;Genetics:
A;Gene: ATSP:F28P10.190
A;Map position: 3
A;Introns: 63/2; 159/1; 230/3; 238/3; 299/1; 372/2; 417/3; 470/1; 503/1

Query Match   10.4%;   Score 285; DB 2;   Length 571;
Best Local Similarity 23.5%;   Pred. No. 8.5e-16;
Matches   99;   Conservative   84;   Mismatches 164;   Indels   74;   Gaps   15;

Qy   66 EGEPCGD---EGAEAPVEGDIHYQRGSGAPLPPSGSKDQV------GGGGEFGGHDKP 114
               : |:| ||||
Db   104 ESLPCVTKPLLEDEEAP-----KHKLSTHSLLPSKPSSSMVSHDMGISNDSSFG----- 153

Qy   115 KITAWEAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEEN 174
               ||||| ||| |:||
Db   154 ----QAVLNGVNVLCGVGLSTPIAVKSGWGLGLIILPAFGIICFYTGLLRYCL-DSH 207

Qy   175 EDGEVVRVRDSYVAIANACCAPRFPTLGGRVVNVAQIELVMTCILYVVSGNLMYNSFP 234
                | :   : | :|:: ||:
Db   208 PDVQ-----TYPDIGHAA---FGSTGRILVSVILYMELYAMSVEYIILEGDNLSSMFP 257

Qy   235 GLPVS-------QKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCL 286

RESULT 11
JC7961
proton-coupled amino acid transporter 2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Aug-2003 #sequence_revision 25-Aug-2003 #text_change 15-Sep-2003
C;Accession: JC7961
R;Chen, Z.; Kennedy, D.J.; Wake, K.A.; Zhuang, L.; Ganapathy, V.; Thwaites, D.T.
Biochem. Biophys. Res. Commun. 304, 747-754, 2003
A;Title: Structure, tissue expression pattern, and function of the amino acid transport
A;Reference number: JC7961; PMID:12727219
A;Accession: JC7961
A;Molecule type: mRNA
A;Residues: 1-481 <CHB>
A;Experimental source: (Lung)
C;Comment: This transport protein is a second member of H+-coupled, pH-dependent, Na++
sport not only from one cellular compartment to another (across plasma membranes) but a
C;Genetics:
A;Gene: pat2
A;Map position: 10q22
C;Keywords: amino acid transport system; PAT2

Query Match   9.2%; Score 252; DB 2; Length 481;
Best Local Similarity 21.9%; Pred. No. 3.9e-13;
Matches 114; Conservative 89; Mismatches 203; Indels 114; Gaps 22;

Qy   60 MDILKAEGEPCCDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQ-------VGGGGEFGGHDK 113
                              ::|: ::.
Db   1 MSVTKSAGSP---------QVAATVKLDL------VSFPESAKKVQSQDNPVNGSSSSESEKT 49

Qy   114 PKITAWEAGWNVTNAIQGMFVLGLPYAILHGGYL-GLFLIIFAAVVCCYTGKILIAC--- 169
                                                    ::
Db   50 KGITGFQTLVHVKGNMGTGIGLPLAVKNAGILMGPLSLLVMGLIACHCMHILVRCAQR 109

Qy   170 -LYEENED---------GEVV--------------RVRDSYVAIANA---CCAPRPPTLG 202
                                   :::
Db   110 FCHRLNKPFMDYGDTVMHGLASSPNTWLQSHAHWGRHAVSFFLIVTQLGFCCV------- 162

Qy   203 GRVV----NVAQIIELV--------MTCILYVVVSGNLMYNSFPGLPVSQKSWSIIAT 248
                            ::|::|
Db   163 -YIVFLADNLKQVEAVNSTTISCHKNETVVLTPTIDSRLYMLAF--LP-------- 208

Qy   249 AVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFYIDVKKFPI 308
                                                :
Db   209 -VLGLLVFIRNLRVLTIFSLLANVSMLVSLVIIGGQYIIQGIPDPS--QLPIVASWKTYPL 265

Qy   309 SIGIIVFSYTS-QIFLPSLEGNMQQPSEFHCMNNWTHIAACVLKGLFALV--AYLTWAD 364
                 :   :::
Db   266 FFGTAIFSFESIGVVLP-LENKMKDARRFPTILS---LGMSIITTLYIAIGALGYLRFGD 321

Qy   365 ETKEVITDNLPGS-IRAVVNIFLVAKALLSYPLPFFAAEVLEKSLFQEGSRAFFPACYS 423
                                    :::
Db   322 DIKASITLNLPNCWLYQSVKLLYVVGILCTHALQFYVPAEIIIPLAVSQVS------- 372

Qy   424 GDGRLKSWG-----LTLRCALVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLR 479
                :::
Db   373 -----KRWALPVDLSIRLALVCVTCMLALLIPRLDLVLSLVGSVSSALALLIPPLLEVT 427

Qy   480 LLWRK-LLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEA 518
                  :::
Db   428 TYYGEGMSPLTITKDALISILGFMGFVVGTYQALDELIRS 467

RESULT 12
S55188
hypothetical protein YJR001w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein J1409; hypothetical protein YJR83.4
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S55188; S57016
R;de Haan, M.; Smits, P.H.M.; Grivell, L.A.
submitted to the EMBL Data Library, May 1995
A;Reference number: S55183

Db   258 NASLSIGGFHLDAPRLFALLTTLAVLPTWLRDLSVLSYIS-----AGGVIASVLVVLCL 312
                |||: :   :::| |||
Qy   287 SRARDWAW--EKVKFY-----IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCM 339

Db   313 F----WVGLLVDDVGIHSKGTPLNLATLPVSVGLYGCYSGHGVFPNIYTSMAKPSQFSAV 368
                      : ::   :.:::: ::.::| | :::| :::::| :::
Qy   340 MNWTHIAACVLKGLFALVAYLTWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFF 399

Db   369 LLASFGICTLMYAGVAVNGYSMPGESTESQFTLNLPQDLVA-SKIALWTTKESTYALTLS 427
                |:||||:   :   :: | :|::::::|
Qy   400 AAEVLEKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVFTLLMAIYVPHFALLMGL 459

Db   428 PVAMSLEE-------LIPSNY--GKSRFYAIAIRSALAISTLLVGLAIPFFVNLLSL 475
                |: ||:| |:||| ||:   :|::|::   : ::: ::|
Qy   460 T 460

Db   476 T 476

RESULT 10
T51506
hypothetical protein F5E19_80 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51506
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51506
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-426 <SAT>
A;Cross-references: UNIPROT:Q91FE3; EMBL:AL391147
A;Experimental source: cultivar Columbia; BAC clone F5E19
C;Genetics:
A;Map position: 5
A;Introns: 48/1; 247/3
A;Note: F5E19_80
C;Superfamily: Arabidopsis amino acid transport protein I

Query Match   9.9%; Score 272; DB 2; Length 426;
Best Local Similarity 24.3%; Pred. No. 7.3e-15;
Matches 100; Conservative 75; Mismatches 185; Indels 52; Gaps 10;

Qy   124 NVTNAIQGMFVLGLPVAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVVRVR 183
               ||| |||:| ||| | :|| ::| | |::
Db   41 NMVGMLIGLGQLSMPYAVESGGWMSIFLLISFGILTTYTSHILGKCIRRNPKS------- 93

Qy   184 DSYVAIANACCAPRPPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPG------L 236
               ||| :   :::   ::::
Db   94 KSYSDIGYSA----FGRHGRLIVCLFIYLEIFMALVSYTISLHDNISAAFPATFSNHGHF 149

Qy   237 PVSQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEK 296
               |:|:|||||:|| :| ::|:
Db   150 PAAK--LTAVAVAIALPSLWIRDLSSISFSLSSGGILMSAIIFGSVVTAIFGG------- 200

Qy   297 VKFYIDVKKFPI------SIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNNWTHIAAC 348
               ||||||:| |||||||||:| ::|::|
Db   201 ---VIDDGKIPVLRLENIPTVSGIYLFSFGGHIVFPNLYTSMKDPSKFTKVSIVSFATVT 257

Qy   349 VLKGLFALVAYLTWADETKEVITDNLPGSIRAVVNIFLVAKAL--LSYPLPFFAAEVL 405
               |:||| ||| ||::|||:| |||:| :||:|
Db   258 ALYGALAITGAKMFGPSVNSQTTLSLPKHL-VVTKIALWATVLTPMTKYALEFAPLAIQL 316

Qy   406 EKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVFTLLMAIYVPHFALLMGLTGSLTG 465
               ||:|| ::| | |||: |:||||:||:||||| | |::| ||| :|||||:||:|
Db   317 ERSL-------PSTMTDRTKLVARGL-MGSALLLVILALITVPYFGVVLSLTGSLVS 366

Qy   466 AGLCFLLPSLPHLRLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIE 517
               ||: || |:: ||||| |||||| ||||||||||| ::
Db   367 VTIAVTLPSAFYLKICWDGMTKFTRAANLGFVLGCVLGVLGSFSSSKLLVK 418

```
A;Accession: S55188
A;Molecule type: DNA
A;Residues: 1-602 <DEH>
A;Cross-references: UNIPROT:P47082; EMBL:X87611; NID:g854567; PIDN:CAA60922.1; PID:g8545
R;de Haan, M.; Grivell, L.A.; Smits, P.H.M.
   submitted to the Protein Sequence Database, September 1995
A;Reference number: S56771
A;Accession: S57016
A;Molecule type: DNA
A;Residues: 1-602 <ZAG>
A;Cross-references: EMBL:Z49501; NID:g1015619; PIDN:CAA89523.1; PID:g1015620; MIPS:YJR0
C;Genetics:
A;Cross-references: SGD:S0003761
A;Map position: 10R


Query Match
Best Local Similarity    8.8%;    Score 241; DB 2; Length 602;
                        24.2%;    Pred. No. 4.2e-12;
Matches 114; Conservative  85; Mismatches 179; Indels 94; Gaps 22;

Qy    89 SGAPLPPSGSGKDQV----GGGGEFGHDKPKI----------------TAWEAGWNV 125
                                                          ||||||||
Db   158 SGSILRPTASLTEMSGGAGRRFTNNDMDSIVKRVEGVDGKVTLLAGQSTAPQTIFNS 217

Qy   126 TNAIQGMFVLGLPYAILHGGY-LGL-FLIIFAAVVCCYTGKILIACLYEENEDGEVVRVR 183
         ::||:|    | | |     | :||  : :|||:
Db   218 INVLIGLGLALALPLGLKYAGWVIGLTMLAIFALATFC-TAELLSRCL--DTDPTLI-- 270

Qy   184 DSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVSGNLMYNSFPGLPVSQKSW 243
                          |:     :|:     : :| :||
Db   271 -SYADLGYAA-----FGTKGRALISALFTLDLLGSGVSLVILFGDSLNALFP--QYSTTFF 323

Qy   244 SIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYC----------LSRARDW 292
                   | :|    |  :::  :|:  : :|:      ::           |:: ||
Db   324 KIVSFFIVTPPVFIP-LSVLSNISLLGILS-TTGTVLVICCCGLYKSSSPGSLVNPMETS 381

Qy   293 AWEKVKFYIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKG 352
          :    :  |:    :|  : : :  :  :  |     |   :
Db   382 MWP-----IDLKFHLCLSIGLLSACWGGHAVFPNLKTDMRHPDKFKDCLKTTYKITSVTDI 436

Qy   353 LFALVALVLTWADETKEVITDN--------LPGSIRAVVNIFLVAKALLSYPLPFFAAV 402
         : ::|  :|    :    | :         :|  |     :::|
Db   437 GTAVIGFLMFGNLVKDEITKNVLLTEGYPKFVYGLISALMTIIPIAKTPLN-ARPIVSVL 495

Qy   403 EVL-EKSLFQEGSRAFFPACYSGDGRLKSWGLTL--RCALVVFTLLMAIYVPHFALLMGL 459
          :|       :  : |                    ::   :::|:|  | || ::
Db   496 DVLMNVQHIDEAASAI-------KRRAAKGLQVFNRIFINVVFVLIAINFPEFDKII-- 545

Qy   460 TGSLTGAGLCF----LLPSLFHLRLL------WRKLLWH-QVFFDVAIFVIG 500
          | :|:|||||    ||| ::  ||         |  :     : :  :|
Db   546 --AFLGAGLCFTICLLPCWFYLRLCKTTIKPWERVACHVTICISVVLSTLG 595


RESULT 13
T49959
hypothetical protein F8M21.130 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49959
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
   submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24493
A;Accession: T49959
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-423 <BEV>
A;Cross-references: UNIPROT:Q9LXF8; EMBL:AL353993; GSPDB:GN00063; ATSP:F8M21.130
A;Experimental source: cultivar Columbia; BAC clone F8M21
C;Genetics:
A;Gene: ATSP:F8M21.130
A;Map position: 5
A;Introns: 48/1; 251/3
C;Superfamily: Arabidopsis amino acid transport protein I
```

```
Query Match
Best Local Similarity    8.6%;    Score 236; DB 2; Length 423;
                        23.9%;    Pred. No. 7.2e-12;
Matches  99; Conservative  82; Mismatches 172; Indels 62; Gaps 17;

Qy   127 NAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVVRVRDSY 186
         ||:  : :: :||| | ||  : |:: :  | :
Db    44 NALSGVGILSVPYALASGGWLSLILLFTVAITTFYCALLIKRCM------EMDPLLRSY 96

Qy   187 VAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVSG--NLMYN----SFPGLPVS 239
         | :    :       ||:     :::  :
Db    97 PDIGYKA-----FGNTGRVIVSIFMNLEIYLVATSFIIEGDNLNKLFSNVGLNFMGLEFQ 152

Qy   240 QKSWSIIATA-VLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVK 298
          |  ||:   |:    :  :::       :||||  :|::
Db   153 GKQMFIMVALIILPSVWLDNMRILSYVSASGVPASGVILASIFSV------GAFEGVG 205

Qy   299 F-YIDVKKFPI----SIGIIVFSYTSQIFLPSLEGNMQQPSFH----CMMNWTHIAA 347
          : |:||   :    ||||:|:||||||||||||||:  :      |::|| ||
Db   206 FKNNDSEVFRLNGVATSVSLYAFCYCAHPVFPTLYTSMKNKRQFSNVMIICTICTFIYA 265

Qy   348 CVLKGLFALVAYLTWADETKEVITDNLP-GSIRAVVNIFLVAKALLSYPLPFFA-AVEVL 405
          ::|  |  :|    ::: :   : :: :||| |::: : |  ||   :
Db   266 SV-----AVLGYLMYGSDVESQITLNLFTDKLSSKVAIW----TTLVNPIAKFALMVTPI 316

Qy   406 EKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTG 465
         |||||||| |||   ||||||||||: :::|||| |||:|| :| ||||||| |  |
Db   317 IDAMRSRFSRVL------PNKRASGFLLSTILIVTSNVIVALLLPFFGDLMSLVGAFLS 368

Qy   466 AGLCFLLPSLFHLRLLWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYR 520
         : :  |:|| :|       :::::    :  |  :   ||  :
Db   369 ASASVILPCLCYLKISGK--YQRLGFETLVLI--GI-TLTGIVVVITGTYQAVK 417


RESULT 14
T34016
hypothetical protein Y4C6B.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34016
R;Layman, D.; Graves, T.; Yoakum, M.
   submitted to the EMBL Data Library, February 1999
A;Description: The sequence of C. elegans cosmid Y4C6B.
A;Reference number: Z21460
A;Accession: T34016
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-484 <LAY>
A;Cross-references: UNIPROT:Q9UA29; EMBL:AF125971; PIDN:AAD14765.1; GSPDB:GN00022; CESP:
A;Experimental source: strain Bristol N2; clone Y4C6B
C;Genetics:
A;Gene: CESP:Y4C6B.2
A;Map position: 4
A;Introns: 97/2; 272/3; 312/1; 403/1; 450/2
C;Superfamily: Arabidopsis amino acid transport protein I


Query Match
Best Local Similarity    8.1%;    Score 223; DB 2; Length 484;
                        22.5%;    Pred. No. 1e-10;
Matches 101; Conservative  91; Mismatches 179; Indels 78; Gaps 20;

Qy   112 DKPKITAWEAGWNVTNAIQGMFVLG--------LPYAILH-GGYLGLFLIIFAAVVCCY 161
         | |    :  || : :|||  ||:|         ||| || |||||||||||||||||
Db     2 DQPHIN--EKGMSI--LVTALFIVGETAGGLIALPTAIVSTGAVTGAFLLLVAAVVCTY 57

Qy   162 TGKIL-----IACLYEENEDGEVVRVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELV 215
         ||:|     :||| |            : |  |
Db    58 TGILLAENWTILQELYPEYRD----HCRKYPAMGLRAIGPKFAHFVSAMLQVTQFGTAV 113

Qy   216 MTCILYVVVSGNLMVNSFPGLPVSQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHF 275
         :|   |::    ::|    :  ::  ||| :
Db   114 VFVLLAAKNGNMLHANF-GTHVSFCYMILIVGLLVFPFTLPKSPKDF-WYAVVAAMIST 171

Qy   276 VINILVIAYCLSRARDW-AWEKVKFY--IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQ 332
         :|: |:::|| |||:|  |:|: :    |:::::| ::|::::::||::
```

```
Db    172  TISVVLIIF--GTVQDYEVCHKEVFYPSFNFSKTLMSFGTIMFSYGGHCAPPTIQHDMKK 229
                 :   ::  :   :    :: :::.:     ::   .:.
Qy    333  PSEFHCMMNWTHIAACVLKGLF----ALVAYLTWADETKEVITDNLPG-SIRAVVNIFLV 387
                    ||    |  |:  :  ::. :  : . ::  |:|||:
Db    230  PHHF----SRSIIIAFIIIFIPIFYLPVSMSGYFVYGSSLTDSIIPSIQNINIQTIVNILIS 285

Qy    388  AKALLSYPLPFFAAVELEKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVVFTLL-M 446
              |:.||   |:| : :::.          |:   ||: :: :::
Db    286  LHVSLALTITFNPLNQBFEEIL--NMSHDF--------GWQRIVARALVMLSVVFV 331

Qy    447  AIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLW--------RKLLWHQVFFD-- 493
              ||:.|:.| : |  | .   |  ::           |:||:
Db    332  AESVPNFGVLLDLVGGSTITMALVVFIVFNLALTTIKKKKENKDNEEKLTIKEIFQNSD 391

Qy    494  --------VAIFVI-GGICSVSGFVHSL 512
                       ||:|| :: ::
Db    392  KIKLLANLILVPAIFGGVAATTSAMQTM 420
```

RESULT 15
JC7626
amino acid transport system N transporter, SN2 - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C;Accession: JC7626
R;Nakanishi, T.; Sugawara, M.; Huang, W.; Martindale, R.G.; Leibach, F.H.; Ganapathy, M.
Biochem. Biophys. Res. Commun. 281, 1343-1348, 2001
A;Title: Structure, function, and tissue expression pattern of human SN2, a subtype of t
A;Reference number: JC7626; MUID:21139776; PMID:11243884
A;Contents: Liver cell line HepG2
A;Accession: JC7626
A;Molecule type: mRNA
A;Residues: 1-472 <NAK>
A;Cross-references: UNIPROT:Q96FQ4; GB:AF276889; GB:AF196972
C;Comment: This SN2, an amino acid transporter, mediates Na+-coupled transport of system
ine, and glycine. The transport function of SN2 is pH-dependent and Li+-tolerant.
C;Genetics:
A;Gene: sn2
A;Map position: X p11.23
C;Keywords: amino acid transport; transport system

```
Query Match     8.1%;   Score 222.5;  DB 2;  Length 472;
Best Local Similarity  22.0%;   Pred. No. 1.1e-10;
Matches   93;  Conservative  79;  Mismatches  152;   Indels  99;  Gaps  16;

Qy    123  WNVTNAIQGMFVLGLPYAILHGG--YLGLFIIFAAVVCCYTGKILIAC------LYE 172
              :|.:||| .:||| |:.:::      :::   ||::    ::  : ::.  :
Db     55  FNLSNAIMGSGILGLAYAMAHTGVIFLALLLCI--ALLSSYSIHLLLTCAGIAGIRAYE 112

Qy    173  E------NEDGEVVRVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILVV---- 222
                         : |   :     : |  ||| :|:.  |  ||:|:
Db    113  QLGQRAFGPAGKVV-------VATVICLHNVGAMSYLPIIKSELPLVIGTFLYMDPEG 164

Qy    223  --VVSGNLMYNSFPGLPVSQKSWSIIATAVLLPCAFLKNLKAV---SKFSLLCTLAHFVI 277
                  ||:|  ||  |:: |   ::: : ||   |   ::|:|:|     |||:|: :::
Db    165  DWFLKGNLLI-----------IIVSVLILPLALMKHLGYLGYTSGLSLTCML-FFLV 210

Qy    278  NILVIAYCL---------------SRARDWAWEKVKFYIDVKKPPISIGIIVFSY 317
              :::                        : :  :::  ::  :: ::::
Db    211  SVIYKKFQLGCAIGHNETAMESEALVGLPSQGLNSSCEAQMFTVD-SQMSYTVPIMAFAF 269

Qy    318  TSQIFLPSLEGNMQQPSE--FHCMMNWTHIAACVLKGLFALVAYLTWADETKEVITD--N 373
              ::  :  :|  |:       :: :   |:|||   | |: :  .|| |:: |::|::
Db    270  VCHPEVLPIYTELCRFPSKRRMQAVANVSIGAMFCMYGLTATFGYLTFYSSVKAEMLHMYS 329

Qy    374  LPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGL 433
              | | :: :  | :: |:
Db    330  QKDPLLICVRIAVLLAVTLTVPVVLFPIRRALQQLLP------PGKAFSWPR 375

Qy    434  TLRCALVVFTL--LMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLL--------W 482
              ::    ::: : : :: |  :    :|:: :  |: : : :::::::::
Db    376  HVAIALLLLVLVNVLVICVPPIRDIFGVIGSTSAPSLIFILPSIFYLRIVPSEVEPFLSW 435
```

This Page Blank (uspto)

```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     November  8, 2004, 18:40:25 ; Search time 464 Seconds
                                 (without alignments)
                                 651.017 Million cell updates/sec

Title:       US-09-940-919-2
Perfect score: 2750
Sequence:    1 MATLLRSKLSNVATSVSNKS.........SGFVHSLEGLIBAYRTNAED 525

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:    1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters:     1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   UniProt_02:*

             1: uniprot_sprot:*
             2: uniprot_trembl:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2750	100.0	525	1	VIAA_HUMAN	Q9h598 homo sapien
2	2731	99.3	525	1	VIAA_MACFA	Q95ke2 macaca fasc
3	2703	98.3	525	1	VIAA_MOUSE	O35633 mus musculu
4	2700	98.2	525	1	VIAA_RAT	O35458 rattus norv
5	2434.5	88.5	518	2	Q6PF45	Q6pf45 xenopus lae
6	2434.5	88.5	518	2	AAH57733	Aah57733 xenopus l
7	2422.5	88.1	518	2	Q6DIV6	Q6div6 xenopus tro
8	1283.5	46.7	638	2	Q75RZ6	Q75rz6 ciona intes
9	1283.5	46.7	638	2	BAD06308	Bad06308 ciona int
10	1081.5	39.3	549	2	Q7QHB6	Q7qhb6 anopheles g
11	1075	39.1	549	2	Q9V6Y7	Q9v6y7 drosophila
12	858	31.2	486	1	UN47_CAEEL	Un47 caenorhabdi
13	853	31.0	164	2	Q8SPU0	Q8spu0 macaca mula
14	375.5	13.7	455	2	Q7XUV8	Q7xuv8 oryza sativ
15	367.5	13.4	571	2	Q6YU97	Q6yu97 oryza sativ
16	367.5	13.4	571	2	BAD08181	Bad08181 oryza sat
17	346.5	12.6	526	2	Q8LPF4	Q8lpf4 arabidopsis
18	341	12.4	424	2	Q94D27	Q94d27 oryza sativ
19	334.5	12.2	550	2	Q8GYS4	Q8gys4 arabidopsis
20	332	12.1	550	2	Q949Y8	Q949y8 arabidopsis
21	320.5	11.7	516	2	Q9LZL5	Q9lzl5 arabidopsis
22	320.5	11.7	536	2	O80668	O80668 arabidopsis
23	317	11.5	543	2	Q9LZL4	Q9lzl4 arabidopsis
24	316.5	11.5	407	2	Q8H139	Q8h139 arabidopsis
25	316	11.5	425	2	Q7XQJ4	Q7xqj4 oryza sativ
26	310	11.3	397	2	Q7XUV9	Q7xuv9 oryza sativ
27	306	11.1	449	2	O17275	O17275 caenorhabdi
28	303	11.0	424	2	Q84YL8	Q84yl8 oryza sativ
29	302	11.0	491	2	Q7Q4M8	Q7q4m8 anopheles g
30	301.5	11.0	478	2	Q9SR29	Q9sr29 arabidopsis
31	299.5	10.9	455	2	O45936	O45936 caenorhabdi
32	299	10.9	413	2	Q8GXJ5	Q8gxj5 arabidopsis
33	296.5	10.8	432	2	Q9SF09	Q9sf09 arabidopsis
34	295.5	10.7	436	2	Q9SVG0	Q9svg0 arabidopsis
35	295	10.7	422	2	Q6ESP5	Q6esp5 oryza sativ
36	295	10.7	429	2	Q9MBG9	Q9mbg9 arabidopsis
37	294.5	10.7	436	2	Q8LAI4	Q8lai4 arabidopsis
38	285	10.4	571	2	Q9SV32	Q9sv32 arabidopsis
39	284.5	10.3	476	2	Q7PM43	Q7pm43 anopheles g
40	284	10.3	427	2	Q9FKY3	Q9fky3 arabidopsis
41	283.5	10.3	443	2	Q8LQU5	Q8lqu5 oryza sativ
42	279.5	10.2	418	2	Q84TX7	Q84tx7 oryza sativ
43	275.5	10.0	432	2	O24406	O24406 arabidopsis
44	272	9.9	426	2	Q9LFE3	Q9lfe3 arabidopsis
45	268.5	9.8	481	2	Q9SR30	Q9sr30 arabidopsis

```
                              ALIGNMENTS

RESULT 1
VIAA_HUMAN
ID   VIAA_HUMAN          STANDARD;      PRT;    525 AA.
AC   Q9H598; Q8N489;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Vesicular inhibitory amino acid transporter (GABA and glycine
DE   transporter) (Vesicular GABA transporter) (hVIAAT).
GN   Name=VIAAT; Synonyms=VGAT;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22007589; PubMed=12031963;
RA   Chessler S.D., Simonson W.T., Sweet I.R., Hammerle L.P.;
RT   "Expression of the vesicular inhibitory amino acid transporter in
RT   pancreatic islet cells: distribution of the transporter within rat
RT   islets.";
RL   Diabetes 51:1763-1771(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   PubMed=14702039; DOI=10.1038/ng1285;
RA   Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA   Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA   Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA   Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA   Nagahari K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA   Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA   Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA   Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA   Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA   Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA   Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA   Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA   Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA   Imose N., Musashino K., Yuuki F., Oshima A., Sasaki N., Aotsuka S.,
RA   Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA   Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA   Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA   Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA   Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA   Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA   Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikena Y., Okamoto S.,
RA   Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA   Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA   Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA   Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA   Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
```

RA Nakai K., Yada T., Nakamura Y., Chara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052; DOI=10.1038/414865a;
RA Deloukas P., Matthews L.H., Ashurst J.L., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehvaslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams S.A., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22382257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP TISSUE SPECIFICITY.
RX MEDLINE=22110478; PubMed=12115694; DOI=10.1002/cne.10272;
RA Jellali A., Stussi-Garaud C., Gasnier B., Rendon A., Sahel J.-A.,
RA Dreyfus H., Picaud S.;
RT "Cellular localization of the vesicular inhibitory amino acid
RT transporter in the mouse and human retina.";
RL J. Comp. Neurol. 449:76-87(2002).
CC -!- FUNCTION: Involved in the uptake of GABA and glycine into the
CC   synaptic vesicles.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Intracellular
CC   membrane vesicles (By similarity).
CC -!- TISSUE SPECIFICITY: Retina. Expressed throughout the horizontal
CC   cells or more specifically at the terminals.
CC -!- SIMILARITY: Belongs to the amino acid/polyamine transporter family

DR EMBL; AY044836; AAK98782.1; -.
DR EMBL; AK055051; BAB70846.1; -.
DR EMBL; AL133519; CAC15529.2; -.
DR EMBL; BC036458; AAH36458.2; ALT_INIT.
DR EMBL; BC053582; AAH53582.1; -.
DR Genew; HGNC:15800; VIAAT.
DR InterPro; IPR002422; AA/rel_permease2.
DR Pfam; PF01490; Aa_trans; 1.
KW Glycoprotein; Neurotransmitter transport; Transmembrane; Transport.
FT DOMAIN   1   133 Cytoplasmic (Potential).
FT TRANSMEM 134   154 Potential.
FT DOMAIN   155   204 Vesicular lumen (Potential).
FT TRANSMEM 205   225 Potential.
FT DOMAIN   226   242 Cytoplasmic (Potential).
FT TRANSMEM 243   263 Potential.
FT DOMAIN   264   265 Vesicular lumen (Potential).
FT TRANSMEM 266   286 Potential.
FT DOMAIN   287   305 Cytoplasmic (Potential).
FT TRANSMEM 306   326 Potential.
FT DOMAIN   327   341 Vesicular lumen (Potential).
FT TRANSMEM 342   362 Potential.
FT DOMAIN   363   383 Cytoplasmic (Potential).
FT TRANSMEM 384   404 Potential.
FT DOMAIN   405   438 Vesicular lumen (Potential).
FT TRANSMEM 439   459 Potential.
FT DOMAIN   460   461 Cytoplasmic (Potential).
FT TRANSMEM 462   482 Potential.
FT DOMAIN   483   489 Vesicular lumen (Potential).
FT TRANSMEM 490   510 Potential.
FT DOMAIN   511   525 Cytoplasmic (Potential).
FT CARBOHYD 341   341 N-linked (GlcNAc...) (Potential).
FT CONFLICT 261   261 K -> R (in Ref. 4; AAH36458).
SQ SEQUENCE 525 AA; 57415 MW; C3C4D07B6B8DE745 CRC64;

Query Match   100.0%; Score 2750; DB 1; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.7e-195;
Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEBAVGPAHCDDLDFEHRQGLQM   60
         |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEVGPAHCDDLDFEHRQGLQM   60

Qy   61 DILKAEGSPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE   120
         ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE   120

Qy 121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180

Qy 181 RVRDSYVAIANACCAPRFPTLGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 RVRDSYVAIANACCAPRFPTLGGRVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ   240

Qy 241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300

Qy 301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMWTHIAACVLKGLFALVAYL 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMWTHIAACVLKGLFALVAYL 360

Qy 361 TWADETKEVITDNLPGSIRAVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420

```
FT   TRANSMEM      490    510       Potential.
FT   DOMAIN        511    525       Cytoplasmic (Potential).
FT   CARBOHYD      341    341       N-linked (GlcNAc...) (Potential).
SQ   SEQUENCE   525 AA;  57393 MW;  2CF17504CB2FD1C9 CRC64;

Query Match                                        Score 2731;  DB 1;  Length 525;
Best Local Similarity   99.3%;                     Pred. No. 6.9e-194;
Matches  521;  Conservative  1;  Mismatches  3;  Indels  0;  Gaps  0;

Qy    1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM   60
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    1 MATLLRSKLSNVATSVSNKSQAKVSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM   60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSSKDQVGGGEFGGGHDKPKITAWE  120
          |||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||||||
Db   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGAGEFGGGHDKPKITAWE  120

Qy  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV  180

Qy  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIELVMTCILYVVVSGNLMYNSFPGLPVSQ  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRDSYVAIANACCAPRFPTLGGRVVNVAQIELVMTCILYVVVSGNLMYNSFPGLPVSQ  240

Qy  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRADWAWEKVKFY  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRADWAWEKVKFY  300

Qy  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHAACVLKGLFALVAYL  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHAACVLKGLFALVAYL  360

Qy  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA  420

Qy  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL  480

Qy  481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
          |||||||||||||||||||||||||||||||||||||||||||||
Db  481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
```

RESULT 3
VIAA_MOUSE
```
ID   VIAA_MOUSE     STANDARD;      PRT;   525 AA.
AC   O35673;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Vesicular inhibitory amino acid transporter (GABA and glycine
DE   transporter) (Vesicular GABA transporter) (mVGAT).
GN   Name=Viaat; Synonyms=Vgat;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=BALB/c; TISSUE=Brain;
RX   MEDLINE=98055669; PubMed=9395291;
RA   Sagne C., El Mestikawy S., Isambert M.-F., Hamon M., Henry J.-P.,
RA   Giros B.P., Gasnier B.;
RT   "Cloning of a functional vesicular GABA and glycine transporter by
RT   screening of genome databases.";
RL   FEBS Lett. 417:177-183(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX   MEDLINE=22461323; PubMed=12573541;
RA   Ebihara S., Obata K., Yanagawa Y.;
```

```
Db  361 TWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL  480
Qy  481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
          |||||||||||||||||||||||||||||||||||||||||||||
Db  481 LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
```

RESULT 2
VIAA_MACFA
```
ID   VIAA_MACFA     STANDARD;      PRT;   525 AA.
AC   Q95KE2;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Vesicular inhibitory amino acid transporter (GABA and glycine
DE   transporter) (Vesicular GABA transporter) (QccE-21148).
GN   Name=VIAAT; Synonyms=VGAT;
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Involved in the uptake of GABA and glycine into the
CC       synaptic vesicles (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Intracellular
CC       membrane vesicles (By similarity).
CC   -!- SIMILARITY: Belongs to the amino acid/polyamine transporter family
CC       II.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AB062931; BAB60726.1; -.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; AA_trans; 1.
KW   Glycoprotein; Neurotransmitter transport; Transmembrane; Transport.
FT   DOMAIN          1    133       Cytoplasmic (Potential).
FT   TRANSMEM      134    154       Potential.
FT   DOMAIN        155    204       Vesicular lumen (Potential).
FT   TRANSMEM      205    225       Potential.
FT   DOMAIN        226    242       Cytoplasmic (Potential).
FT   TRANSMEM      243    263       Potential.
FT   DOMAIN        264    265       Vesicular lumen (Potential).
FT   TRANSMEM      266    286       Potential.
FT   DOMAIN        287    305       Cytoplasmic (Potential).
FT   TRANSMEM      306    326       Potential.
FT   DOMAIN        327    341       Vesicular lumen (Potential).
FT   TRANSMEM      342    362       Potential.
FT   DOMAIN        363    383       Cytoplasmic (Potential).
FT   TRANSMEM      384    404       Potential.
FT   DOMAIN        405    438       Vesicular lumen (Potential).
FT   TRANSMEM      439    459       Potential.
FT   DOMAIN        460    461       Cytoplasmic (Potential).
FT   TRANSMEM      462    482       Potential.
FT   DOMAIN        483    489       Vesicular lumen (Potential).
```

```
RT    "Mouse vesicular GABA transporter gene: genomic organization,
RT    transcriptional regulation and chromosomal localization.";
RL    Brain Res. Mol. Brain Res. 110:126-139(2003).
RN    [3]
RP    SEQUENCE FROM N.A. (ISOFORM 1).
RC    STRAIN=C57BL/6; TISSUE=Brain;
RX    MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA    Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA    Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA    Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA    Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA    Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA    Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA    Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA    Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA    Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA    Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA    Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA    Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA    Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA    Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA    Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA    Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA    Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT    "Generation and initial analysis of more than 15,000 full-length human
RT    and mouse cDNA sequences.";
RL    Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN    [4]
RP    TISSUE SPECIFICITY.
RX    MEDLINE=22110478; PubMed=12115694; DOI=10.1002/cne.10272;
RA    Jellali A., Stussi-Garaud C., Gasnier B., Rendon A., Sahel J.-A.,
RA    Dreyfus H., Picaud S.;
RT    "Cellular localization of the vesicular inhibitory amino acid
RT    transporter in the mouse and human retina.";
RL    J. Comp. Neurol. 449:76-87(2002).
CC    -!- FUNCTION: Involved in the uptake of GABA and glycine into the
CC        synaptic vesicles.
CC    -!- SUBCELLULAR LOCATION: Integral membrane protein. Intracellular
CC        membrane vesicles (By similarity).
CC    -!- ALTERNATIVE PRODUCTS:
CC        Event=Alternative splicing; Named isoforms=2;
CC        Name=1; Synonyms=b;
CC          IsoId=O35633-1; Sequence=Displayed;
CC        Name=2; Synonyms=a;
CC          IsoId=O35633-2; Sequence=VSP_007063;
CC    -!- TISSUE SPECIFICITY: Brain and retina. Localized in horizontal cell
CC        tips at both rod and cone terminals.
CC    -!- SIMILARITY: Belongs to the amino acid/polyamine transporter family
CC        II.
CC    --------------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------------------
DR    EMBL; AB080232; BAC44888.1; -.
DR    EMBL; AB080232; BAC44889.1; -.
DR    EMBL; AJ001598; CAA04864.1; -.
DR    EMBL; BC052020; AAH52020.1; -.
DR    MGD; MGI:1194488; Viaat.
DR    GO; GO:0005887; C:integral to plasma membrane; IC.
DR    GO; GO:0005331; F:gamma-aminobutyric acid transporter activity; IDA.
DR    GO; GO:0005187; F:glycine transporter activity; IDA.
DR    GO; GO:0015816; P:glycine transport; IDA.
DR    InterPro; IPR002422; AA/rel_permease2.
DR    Pfam; PF01490; Aa_trans; 1.
KW    Alternative splicing; Glycoprotein; Neurotransmitter transport;
KW    Transmembrane; Transport.
FT    DOMAIN        1    133       Cytoplasmic (Potential).
FT    TRANSMEM    134    154       Potential.
```

```
FT    DOMAIN      155    204       Vesicular lumen (Potential).
FT    TRANSMEM    205    225       Potential.
FT    DOMAIN      226    242       Cytoplasmic (Potential).
FT    TRANSMEM    243    263       Potential.
FT    DOMAIN      264    265       Vesicular lumen (Potential).
FT    TRANSMEM    266    286       Potential.
FT    DOMAIN      287    305       Cytoplasmic (Potential).
FT    TRANSMEM    306    326       Potential.
FT    DOMAIN      327    341       Vesicular lumen (Potential).
FT    DOMAIN      342    362       Potential.
FT    TRANSMEM    363    384       Cytoplasmic (Potential).
FT    DOMAIN      385    405       Potential.
FT    TRANSMEM    406    438       Vesicular lumen (Potential).
FT    DOMAIN      439    459       Potential.
FT    TRANSMEM    460    461       Cytoplasmic (Potential).
FT    DOMAIN      462    482       Potential.
FT    TRANSMEM    483    489       Vesicular lumen (Potential).
FT    DOMAIN      490    510       Potential.
FT    DOMAIN      511    525       Cytoplasmic (Potential).
FT    CARBOHYD    341    341       N-linked (GlcNAc...) (Potential).
FT    VARSPLIC    515    525       LIEAYRTNAED -> KPAGLET (in isoform 2).
FT                                 /FTId=VSP_007063.
FT    CONFLICT    432    432       G -> E (in Ref. 1).
SQ    SEQUENCE    525 AA;  57381 MW;  EBD63E01A4B54C07 CRC64;
```

```
Query Match             98.3%;  Score 2703;  DB 1;  Length 525;
Best Local Similarity   98.5%;  Pred. No. 8.2e-192;
Matches   518;  Conservative   2;  Mismatches  4;  Indels  2;  Gaps  2;

Qy      1  MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFHRQGLQM   60
           |||||||:|:||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MATLLRSKTNVATSVSNKSQAKVSGMFARMGFQAATDEEAVGFAHCDDLDFHRQGLQM    60

Qy     61  DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQ-VGGGGEFGGHDKPKITAW  119
           |||:|||||||||||||||||||||||  |||||||||||| |||:|||:|||||||||||
Db     61  DILKSEGEPCGDEGAEAPVEGDIHYQRG-GAPLPPSGSKDQAVGAGGEFGGHDKPKITAW  119

Qy    120  EAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYENEDGEV  179
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120  EAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYENEDGEV  179

Qy    180  VRVRDSYVAIANACCAPRPTLGGRVVNVAQIIELVMTCILYVVSGNLMYNSFPGLPVS  239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180  VRVRDSYVAIANACCAPRPTLGGRVVNVAQIIELVMTCILYVVSGNLMYNSFPGLPVS  239

Qy    240  QKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKF  299
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240  QKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKF  299

Qy    300  YIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAY  359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300  YIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVLKGLFALVAY  359

Qy    360  LTWADETKEVITDNLPGSIRAVNIFLVAKALLSYPLFFFAAVEVLEKSLFQEGSRAFFP  419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360  LTWADETKEVITDNLPGSIRAVNLFLVAKALLSYPLFFFAAVEVLEKSLFQEGSRAFFP  419

Qy    420  ACYGGDGRLKSWGLTLRCALIVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLR  479
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420  ACYGGDGRLKSWGLTLRCALIVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLR  479

Qy    480  LLWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    480  LLWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
```

```
RESULT 4
VIAA_RAT
ID    VIAA_RAT     STANDARD;     PRT;   525 AA.
AC    O35458;
DT    10-OCT-2003 (Rel. 42, Created)
DT    10-OCT-2003 (Rel. 42, Last sequence update)
DT    05-JUL-2004 (Rel. 44, Last annotation update)
```

```
Best Local Similarity   98.5%;   Pred. No. 1.4e-191;
Matches 518;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    2;

Qy     1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM  60
         |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MATLLRSKLTNVATSVSNKSQAKVSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM  60

Qy    61 DILKAEGPCGDEGAEAPVEGDIHYQRGSGAPLPPGGSKDQ-VGGGGEFGGHDKPKITAW  119
         |||||| ||||||||||||||||||||| ||||||||||| |||| ||||||||||||||
Db    61 DILKSEGPCGDEGAEPPVEGDIHYQRG-GAPLPSGGSKDQAVGAGEFGGHDKPKITAW  119

Qy   120 EAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEV  179
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   120 EAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEV  179

Qy   180 VRVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVS  239
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   180 VRVRDSYVAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVS  239

Qy   240 QKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKF  299
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240 QKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKF  299

Qy   300 YIDVKKFPISIGIIVSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAY  359
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   300 YIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAY  359

Qy   360 LTWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLPQEGSRAFFP  419
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360 LTWADETKEVITDNLPGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLPQEGSRAFFP  419

Qy   420 ACYSGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLR  479
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   420 ACYGGDGRLKSWGLTLRCALVVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLR  479

Qy   480 LLWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   480 LLWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED  525
```

```
RESULT 5
Q6PF45
ID   Q6PF45        PRELIMINARY;     PRT;    518 AA.
AC   Q6PF45;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC68938 protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
```

```
DE   Vesicular inhibitory amino acid transporter (GABA and glycine
DE   transporter) (Vesicular GABA transporter) (rGVAT).
GN   Name=Viaat; Synonyms=Vgat;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. AND CHARACTERIZATION.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=98007977; PubMed=9349821;
RA   McIntire S.L., Reimer R.J., Schuske K., Edwards R.H., Jorgensen E.M.;
RT   "Identification and characterization of the vesicular GABA
RT   transporter";
RL   Nature 389:870-876 (1997).
RN   [2]
RP   CHARACTERIZATION.
RC   TISSUE=Brain;
RX   MEDLINE=22027589; PubMed=12031963;
RA   Chessler S.D., Simonson W.T., Sweet I.R., Hammerle L.P.;
RT   "Expression of the vesicular inhibitory amino acid transporter in
RT   pancreatic islet cells: distribution of the transporter within rat
RT   islets.";
RL   Diabetes 51:1763-1771(2002).
-!-  FUNCTION: Involved in the uptake of GABA and glycine into the
     synaptic vesicles.
-!-  SUBCELLULAR LOCATION: Integral membrane protein. Intracellular
     membrane vesicles.
-!-  TISSUE SPECIFICITY: Brain. Expressed at high levels within the
     neocortex, hippocampus, cerebellum, striatum, septal nuclei and
     the reticular nucleus of the thalamus. Also expressed in islets
     where it is more abundant in the peripheral/mantle region.
-!-  SIMILARITY: Belongs to the amino acid/polyamine transporter family
     II.
--------------------------------------------------------------------------
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
--------------------------------------------------------------------------
DR   EMBL; AF030253; AAB82950.1; -.
DR   RGD; 621402; Viaat.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
KW   Glycoprotein; Neurotransmitter transport; Transmembrane; Transport.
FT   DOMAIN        1    133       Cytoplasmic (Potential).
FT   TRANSMEM    134    154       Potential.
FT   DOMAIN      155    204       Vesicular lumen (Potential).
FT   TRANSMEM    205    225       Potential.
FT   DOMAIN      226    242       Cytoplasmic (Potential).
FT   TRANSMEM    243    263       Potential.
FT   DOMAIN      264    265       Vesicular lumen (Potential).
FT   TRANSMEM    266    286       Potential.
FT   DOMAIN      287    305       Cytoplasmic (Potential).
FT   TRANSMEM    306    326       Potential.
FT   DOMAIN      327    341       Vesicular lumen (Potential).
FT   TRANSMEM    342    362       Potential.
FT   DOMAIN      363    383       Cytoplasmic (Potential).
FT   TRANSMEM    384    404       Potential.
FT   DOMAIN      405    438       Vesicular lumen (Potential).
FT   TRANSMEM    439    459       Potential.
FT   DOMAIN      460    461       Cytoplasmic (Potential).
FT   TRANSMEM    462    482       Potential.
FT   DOMAIN      483    489       Vesicular lumen (Potential).
FT   TRANSMEM    490    510       Potential.
FT   DOMAIN      511    525       Cytoplasmic (Potential).
FT   CARBOHYD    341    341       N-linked (GlcNAc...) (Potential).
SQ   SEQUENCE   525 AA;  57407 MW;  33C5E8D31B78D510 CRC64;

Query Match             98.2%;  Score 2700;  DB 1;  Length 525;
```

RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC057733; AAH57733.1; -.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
SQ   SEQUENCE 518 AA; 57190 MW; 0DE19ED16BD84COD CRC64;

Query Match   88.5%; Score 2434.5; DB 2; Length 518;
Best Local Similarity   89.5%; Pred. No. 6.3e-172;
Matches 470; Conservative   19; Mismatches   29; Indels   7; Gaps   4;

Qy   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
          |||||||||||||||||||||||||||||||||||||||||:|||||||||| ||||||||
Db   1 MATLIRSKLSNVATSVSNKSQAKVSGMFARMGFQAATDEEALGFAHCDDLDTEHRQGLQM 60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120
          |||:|| :: |||||| ::||||||:   ::|||||||||||||||:|||:||||||||
Db   61 DILKTE-VPTGD---APPEGDIHYQR-DGTGLPPSASKDE-GLCSELSSSEKPQITAWE 113

Qy   121 AGNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   114 AGNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGETV 173

Qy   181 RVRDSYVAIANACCAPRFPTLGGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
          ||||||| ||||||||: ||| |||||||||||||||||||||||||||||||||:||||||
Db   174 RVRDSYVDIANACCAPRFPKLGGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPNLPISQ 233

Qy   241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
          ||||:|||||||||||||||||||||||||| ||||||||||||||||||||||| |||||
Db   234 KSWSIMATAVLLPCAFLKNLKAVSKFSLLCTVAHFVINILVIAYCLSRARDWAWDKVKFY 293

Qy   301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYL 360
          ||||||||||||||||||||||||||||||||| |||||||||||||||||:|||||||
Db   294 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQSPREFHCMNWTHIAACILKGLFALVAYL 353

RESULT 6
AAH57733
ID   AAH57733   PRELIMINARY;   PRT;   518 AA.
AC   AAH57733;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC68938 protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC057733; AAH57733.1; -.
SQ   SEQUENCE 518 AA; 57190 MW; 0DE19ED16BD84COD CRC64;

Query Match   88.5%; Score 2434.5; DB 2; Length 518;
Best Local Similarity   89.5%; Pred. No. 6.3e-172;
Matches 470; Conservative   19; Mismatches   29; Indels   7; Gaps   4;

Qy   1 MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
          |||||||||||||||||||||||||||||||||||||||||:|||||||||| ||||||||
Db   1 MATLIRSKLSNVATSVSNKSQAKVSGMFARMGFQAATDEEALGFAHCDDLDTEHRQGLQM 60

Qy   61 DILKAEGEPCGDEGAEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120
          |||:|| :: |||||| ::||||||:   ::|||||||||||||||:|||:||||||||
Db   61 DILKTE-VPTGD---APPEGDIHYQR-DGTGLPPSASKDE-GLCSELSSSEKPQITAWE 113

Qy   121 AGNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVV 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   114 AGNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGETV 173

Qy   181 RVRDSYVAIANACCAPRFPTLGGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
          ||||||| ||||||||: ||| |||||||||||||||||||||||||||||||||:||||||
Db   174 RVRDSYVDIANACCAPRFPKLGGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPNLPISQ 233

Qy   241 KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFY 300
          ||||:|||||||||||||||||||||||||| ||||||||||||||||||||||| |||||
Db   234 KSWSIMATAVLLPCAFLKNLKAVSKFSLLCTVAHFVINILVIAYCLSRARDWAWDKVKFY 293

Qy   301 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAYL 360
          ||||||||||||||||||||||||||||||||| |||||||||||||||||:|||||||
Db   294 IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQSPREFHCMNWTHIAACILKGLFALVAYL 353

```
Qy   361  TWADETKEVITDNLPGSIRAVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
          ||||||| |||||||||||||||||:||||||||||||||||||||| ||| |||||||
Db   354  TWADETKEVITDNLPSTIRAVVNLFLVSKALLSYPLPFFAAVEVLEKSLFQEGARAFFPN 413

Qy   421  CYSGDGRLKSWGLTLRCALVFFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
          |||||||||||||||||| |||||||||||| ||||||||||||||||||:||||||| |
Db   414  CYGGDGRLKSWGLTLRCALVFFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLKL 473

Qy   481  LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
          ||||| ||||||||:||||| |||||||||||||||||| |  |
Db   474  MWRQLLWHQVFFDVSIFVIGSICSVSGFVHSLEGLIEAYAYNIED 518

RESULT 7
Q6DIV6
ID   Q6DIV6;              PRELIMINARY;       PRT;   518 AA.
AC   Q6DIV6;
DT   01-OCT-2004 (TrEMBLrel. 28, Created)
DT   01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8364;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole body;
RA   Klein S., Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC075429; AAH75429.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE 518 AA;  57155 MW;  1B0B04020851C78B CRC64;

Query Match             88.1%;  Score 2422.5;  DB 2;  Length 518;
Best Local Similarity   89.3%;  Pred. No. 4.9e-171;
Matches  469;  Conservative  16;  Mismatches  33;  Indels  7;  Gaps  4;

Qy     1  MATLLRSKLSNVATSVSNKSQAKMSGMFARMGFQAATDEEAVGFAHCDDLDFEHRQGLQM 60
          |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db     1  MATLLRSKLSNVATSVSNKSQAKVSGMFARMGFQAATDEEALGFAHCDDLDMEHRQGLQM 60

Qy    61  DILKAEGEPCGDGAEAPVEGDIHYQRGSSGAPLPPSGSKDQVGGGEFGGHDKPKITAWE 120
          |||| | | ||| ||::      | |:||::|  |:    |  |     |||||||||||
Db    61  DILKTE-VFSGDPTA---BGDSHYQRDGTGP-PSSASKDE-GLCSELSSYGKPKITAWE 113

Qy   121  AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVCCYTGKILIACLYEENEDGEVV 180
          ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
```

```
Db   114  AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVCCYTGKILIACLYEENEDGETV 173

Qy   181  RVRDSVYAIANACCAPRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQ 240
          ||||||  ||||||||| | |||||||||| ||||||||||||||||:|||||||| |||
Db   174  RVRDSYYDIANACCAPRFPKLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSFPSLPISQ 233

Qy   241  KSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVNILVIAYCLSRARDWAWEKVKFY 300
          |||||||| :||||||||||||||||||||||||||| |||||||||||||||:|||||
Db   234  KSWSIIATAMLLPCAFLKNLKAVSKFSLLCTLAHFVNVLVLVIAYCLSRARDWAWDKVKFY 293

Qy   301  IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEPHCMMNWTHIAACVLKGLFALVAYL 360
          ||||||||||| ||||||||||||||||||||| ||| ||||||||| ||| |||||||
Db   294  IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQSPKEFHCMMNWTHIAACILKGLFALVAYL 353

Qy   361  TWADETKEVITDNLPGSIRAVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPA 420
          ||||||| |||||||||||||||||:||||||||||||||||||||| ||| |||||||
Db   354  TWADETKEVITDNLPSTIRAVVNLFLVSKALLSYPLPFFAAVEVLEKSLFQEGARAFFPN 413

Qy   421  CYSGDGRLKSWGLTLRCALVFFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRL 480
          |||||||||||||||||| |||||||||||| ||||||||||||||||||:||||||| |
Db   414  CYGGDGRLKSWGLTLRCALVFFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLKL 473

Qy   481  LWRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAYRTNAED 525
          ||||| ||||||||:||||| |||||||||||||||||| |  |
Db   474  LWRKLQWHQVFFDVSIFVIGISCSVSGFVHSLEGLIEAFRFNIED 518

RESULT 8
Q75RZ6
ID   Q75RZ6;              PRELIMINARY;       PRT;   638 AA.
AC   Q75RZ6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Vesicular GABA transporter.
GN   Name=Ci-vGAT;
OS   Ciona intestinalis.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC   Phlebobranchia; Cionidae; Ciona.
OX   NCBI_TaxID=7719;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15170699;
RA   Yoshida R., Sakurai S., Horie T., Kawakami I., Tsuda M., Kusakabe T.;
RT   "Identification of neuron-specific promoters in Ciona intestinalis.";
RL   Genesis 39:130-140(2004).
DR   EMBL; AB158401; BAD06308.1; -.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
SQ   SEQUENCE 638 AA;  69532 MW;  707DC701FB409FDA CRC64;

Query Match             46.7%;  Score 1283.5;  DB 2;  Length 638;
Best Local Similarity   44.2%;  Pred. No. 1.5e-86;
Matches  244;  Conservative  88;  Mismatches  117;  Indels  103;  Gaps  6;

Qy    37  TDEEAVGFAHCDDLDFEHRQGLQMDILKAEGEPCGDGAEAPVEGDI-HYQRGSGAPLPP 95
           |:||||||||||||||||||||||| :|: : |      |:||| |:|: |:
Db    18  TSEEKFSFAKSSDQP----------------HPSGSGCQGDGATNPSSNSVSHPERST----- 59

Qy    96  SGSKDQVGGGEFGGHDKPKITAWEAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFA 155
                                   :||:||:|:   |::| |||||||||:||||||
Db    60  -------------------SGVEKPTITAWDAGWNVSNAIQGMFVLGLPYAVLHGGYLGLIIIVT 106

Qy   156  AVVCCYTGKILIACLYEENEDGEVVRDSYVAIANACCAPRFPTLGGRVVNVAQIELV 215
           ||:||||||||||||||||:|||| ||||:||||||| || :| || ||| |:|||
Db   107  AVVCCYTGNILIDCLYETSPSGERLRVRSTYVDLAAHCWGKH---LGGYLVNAAQLIELL 163

Qy   216  MTCILYVVVSGNLMYNSFPGLPVSQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHF 275
           ||:|||||||||||:|||:||| |:||| :||: ||| | |: |||:::|      :|:|
Db   164  MTCVLYVVVSGNLMTNSFPHGPIREAGWSVLACLVLPCIFLRHLRAVSRFSMGCSVAQI 223

Qy   276  VINILVIAYCLSRARDWAWEKVKFYIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSE 335
           |:  || | : | :|| |  | :  :| |  :: :
```

```
Db    344 LPYYQAIEVMEQTMFTGATGWGSLFGTKRHAYGEFTDDTEPIVQSTSFNTDAPSSPSST 403
          ||::||::|||:::|
Qy    411 ----QEG--------------------------SRAFFPACYSGDGRLKSWGLTLRCALVVFTLLMA 447
                                          ::|
Db    404 TDNSEDGLEDKSTTKNVILTLEDDTNKQSSCPSCYSATGDLQVWALVLRAGLVLGTLLMG 463
           :|
Qy    448 IYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLMRKLLWHQVFFDVAIFVIGGICSVSG 507
          :::||||||||||||||||||||||||||::||| ||||||:||||||||| :|:
Db    464 VFIPHFALLMGLTGSLTGSLTGSLAFLFPCAFHLQIKWREIGLDVFIPISGTVCGITG 523
          |||||||| || |:|| :|
Qy    508 FVHSLEGLIEAY 519
          |:||:|||
Db    524 IYFSIQGLYEVY 535
```

RESULT 10
Q7QHB6
```
ID   Q7QHB6       PRELIMINARY;       PRT;    549 AA.
AC   Q7QHB6;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   EbiP8381 (Fragment).
GN   Name=ebiG8381; ORFNames=ENSANGG00000006321;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PEST;
RA   Anopheles Genome Sequencing Consortium;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008816; EAA05252.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR   GO; GO:0006865; P:amino acid transport; IEA.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
FT   NON_TER       549    549
SQ   SEQUENCE   549 AA;  60847 MW;  21A09BD190E02A0 CRC64;
```

```
Query Match             39.3%;
Best Local Similarity   47.1%;         Score 1081.5;  Pred. No. 1.2e-71;  DB 2;  Length 549;
Matches  221;  Conservative  80;  Mismatches 141;  Indels  27;  Gaps  10;
```

```
Qy     72 DEGAEAP--------VEGDIH---YQRGSGAPLPPSGSKDQVGGGGEPGG---HDKP-- 114
Db     81 EEGTDQPGTNQYQETGFNQGDFENGYQAG-GYP-PRQGSVQSFGSDSTFAGGCEGEAPGG 138
Qy    115 -KITAWEAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYE- 172
Db    139 AKINEYQAAWNVTNAIQGMFVLGLPYAVLHGGYWAIIAMVGIAHICCYTGKILVMCLYEP 198
Qy    173 ENEDGEVVRVRDSVVAINACCAPRFPTLGGRVVNVAQIIELVMTCILYVVSGNLMYNS 232
Db    199 DPQTGBPVRVRDSYVSIAKVCFGKK---IGARVVSIAQIIELLMTCILYVVVCCGDLMAGS 255
Qy    233 FPGLPVSQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCLAHFVINILVIAYCLSRARDW 292
Db    256 FPDGALDTRSWMMLCGIFLLPLAFLKSLLHHVSLLSFWCTMAHLLINAIIVGYCLLBIGDW 315
Qy    293 AWEKVKFYIDVKKFPISIGIIVFSYTSQIFLPLSEGNMQQPSEFHCMNWTHIAACVLKG 352
Db    316 GWSKVKWRMDFENFPISLGVIVFSYTSQIFLPTLEGNMEDRSKFNWMLDWSHIAAAAFKA 375
Qy    353 LFALVAYLTWADETKEVITDNL--PGSIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLF 410
Db    376 LFGYICFLTFQNDTQQVITNNLHSPG-FKGLVNFCLVIKAILSYPLPYFAACELLERAFF 434
```

```
Db    224 VVLGITIVYCITRINTWAWSEITISVDMKQFPVSIGVIVFSYTSQIFLPSLEGSMENRGD 283
Qy    336 FHCMNWTHIAACVLKGLFALVAYLTWADETKEVITDNLPPTLRAMINVLLVVKALLSYP 395
Db    284 FRSMLSWSYVASCVTKASFALICFLTWSKDTKDVVTDNLPPTLRAMINVLLVVKALLSYP 343
Qy    396 LPFFAAVEVLEKSLF---------------------------------- 410
Qy    344 LPYYQAIEVMEQTMFTGATGWGSLFGTKRHAYGEFTDDTEPIVQSTSFNTDAPSSPSST 403
Db    411 ---QEG-------------------------------SRAFFPACYSGDGRLKSWGLTLRCALVVFTLLMA 447
Db    404 TDNSEDGLEDKSTTKNVILTLEDDTNKQSSCPSCYSATGDLQVWALVLRAGLVLGTLLMG 463
Qy    448 IYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLMRKLLWHQVFFDVAIFVIGGICSVSG 507
Db    464 VFIPHFALLMGLTGSLTGSLTGSLAFLFPCAFHLQIKWREIGLDVFIPISGTVCGITG 523
Qy    508 FVHSLEGLIEAY 519
Db    524 IYFSIQGLYEVY 535
```

RESULT 9
BAD06308
```
ID   BAD06308     PRELIMINARY;       PRT;    638 AA.
AC   BAD06308;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Vesicular GABA transporter.
GN   CI-VGAT.
OS   Ciona intestinalis.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC   Phlebobranchia; Cionidae; Ciona.
OX   NCBI_TaxID=7719;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yoshida R., Sakurai S., Horie T., Kawakami I., Tsuda M., Kusakabe T.;
RT   "Identification of neuron-specific promoters in Ciona intestinalis.";
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB158401; BAD06308.1; -.
SQ   SEQUENCE   638 AA;  69532 MW;  707DC701F8409FDA CRC64;
```

```
Query Match             46.7%;
Best Local Similarity   44.2%;         Pred. No. 1.5e-86;       Score 1283.5;  DB 2;  Length 638;
Matches  244;  Conservative  88;  Mismatches  117;  Indels  103;  Gaps  6;
```

```
Qy     37 TDEEAVGPAHCDDLDFEHRQGLQMDIKAEGEPCGDEGAEAPVEGDI-HYQRGSGAPLPP 95
Db     18 TSEEKFSFAKSSDQP-----------------HPSGSGCQGDGATNPSSNSVSHPERST------ 59
Qy     96 SGSKDQVGGGGEPGGHDKPKTAWEAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFA 155
Db     60 ---------SGVEKPTITAWDAGWNVSNAIQGMFVLGLPYAVLHGGYLGLILIIVT 106
Qy    156 AVVCCYTGKILIACLYEENEDGEVVRVRDSVVAIANACCAPRFPTLGGRVVNVAQIELV 215
Db    107 AVVCCYTGNILIDCLYETSPSGERLRVRSTYVDLAAHCWGKH--LGGYLVNAAQLIELL 163
Qy    216 MTCILYVVVSGNLMYNSFPGLPVSQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCLAHF 275
Db    164 MTCVLYVVVSGNLMTNSFPHGPIREAGWSVLACLVLFPCIFLRHLRAVSRPSMGCSVAQI 223
Qy    276 VINILVIAYCLSRARDWAWEKVKFYIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSE 335
Db    224 VVLGITIVYCITRINTWAWSEITISVDMKQFPVSIGVIVFSYTSQIFLPSLEGSMENRGD 283
Qy    336 FHCMNWTHIAACVLKGLFALVAYLTWADETKEVITDNLPPTLRAMINVLLVVKALLSYP 395
Db    284 FRSMLSWSYVASCVTKASFALICFLTWSKDTKDVVTDNLPPTLRAMINVLLVVKALLSYP 343
Qy    396 LPFFAAVEVLEKSLF---------------------------------- 410
```

```
Qy  411  QEGSRAFFPACYSGDGRLKSWGLITLRCALVFTLLMAIYVPHFALLMGLTGAGLCF     470
              ||:|  |  | :  |||| |||  |: |  | : ||| |||| |||| ||  | :|
Db  435  RGKPKTMFPVVWELDGELKVWGLAWRLTVLGTIMWAIFPIPHFSILMGFIGSFTGTMLSF  494

Qy  471  LLPSLFHLRLLMRKLLWHQVFFDVAIFVIGGICSVSGFVHSLEGLIEAY           519
              : | |||:|  | || |:   :  : |:|:: : :|||  || |:||:
Db  495  IWPCYFHLKLKGHLLDQKQRAYNYFIIFLGVLFCVVGIYDSGTALIHAF           543


RESULT 11
Q9V6Y7
ID   Q9V6Y7      PRELIMINARY;      PRT;      549 AA.
AC   Q9V6Y7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CG8394-PA.
GN   ORFNames=CG8394;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R.,
RA   Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,

RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FLYBASE;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FLYBASE;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003815; AAF58280.1; -.
DR   FlyBase; FBgn0033911; CG8394.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR   GO; GO:0006865; P:amino acid transport; IEA.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
SQ   SEQUENCE 549 AA; 60838 MW; B0D7A8461FC9F472 CR664;

Query Match             39.1%;   Score 1075;   DB 2;   Length 549;
Best Local Similarity   46.2%;   Pred. No. 3.5e-71;
Matches  217;   Conservative   84;   Mismatches 143;   Indels  26;   Gaps    7;

Qy   66  EGEPCGDEGAE--------APVEGDIHYQRGSG---APLPPSGSK---DQVGGGGEFGGH  111
              |:  |||| | :        ||||||| |||::   ||| |  :      : |||| ||||
Db   84  DGEGGGGDGDGEYRNEYQSTSFNEYDGRYQQTDGFRQGSIASEGSSFVCEGEGGGG-----  138

Qy  112  DKPKITAWEAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIPAAVVCCYTGKILIACLY  171
              ||||||||||||||||||||||| ::  ||||||  ||  : |||||||||||||||||
Db  139  --CKIDEFQAAWNVTNAIQGMFIVSLFPAVLRGGYWAIVAMVGIAHICCYTGKVLVQCLY  196

Qy  172  E-ENEDGEVVRVRDSYVAIANACCAPRFPTLGGRVVNVAQIELVMTCILYVVVSGNLMY   230
              | ||:||||||:| ||::|||  | |||:  ::|::::  ||||:|:||||||| :|::
Db  197  EPDPATGQMVRVRDSYVAIAKVCFGPK---LGRAVSIAQLLMTCILYVVVVCGDLA      253

Qy  231  NSFPGLPVSQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRAR  290
              ||::||| : ||::|: | |: | |  :: :  ::||||  ||||:||::|:::||||:||
Db  254  GTYPQGSFDSRSWMLFVGIFLLPMGFLKSLKMVSTLSFWCTMSHIVINAVILGYCLQIG   313

Qy  291  DWAWEKVKFYIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQQPSEFHCMMNWTHIAACVL  350
              |:|: |  |::::    |:|||:||::||::|||||:||:|:   : |  |:  ||||:|
Db  314  DWGWSKVRFSIDMENFPISLGVIVFSYTSQIFLPTLEGNMDRSKFNWMLDWSHIAAAVF   373

Qy  351  KGLFALVAYLTWADETKEVITDNLPG-SIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSL  409
              |||  |||:|:|:|: ::|:||:||  ||| |:||| :||||||||:|:: ||:|::   :
Db  374  KAGFGYICFLTFQNDTQQVITNNLHSQGFKGMVNFFLVIKALLSYPLPYYAACELLERNF  433

Qy  410  FQEGSRAFFPACYSGDGRLKSWGLTLRCALVFTLLMAIYVPHFALLMGLTGSLTGSAGLC  469
```

```
Db    434  FRGPPKTKPTIWNLDGELKVWGLGFRVGVIVSTLMAIFIPHFSILMGFIGSFTGTMLS 493

Qy    470  FLLPSLFHLRLLWRKLLWHQVFPDVAIPVIGGICSVSGFVHSLEGLIEAY 519
                |::|: |||||  ||::  |::|:  ||:|| ||:||  :|:||
Db    494  FIWPCYFIKIKGHLLDQKEIAKDYLIIGLGVLFGVIGIYDSGNALINAF 543


RESULT 12
UN47_CAEEL
ID   UN47_CAEEL      STANDARD;       PRT;    486 AA.
AC   P34579; O17475;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Vesicular GABA transporter (Uncoordinated protein 47) (Protein unc-
DE   47).
GN   Name=unc-47; ORFNames=T20G5.6;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF GLY-462.
RX   MEDLINE=98007977; PubMed=9349821;
RA   McIntire S.L., Reimer R.J., Schuske K., Edwards R.H., Jorgensen E.M.;
RT   "Identification and characterization of the vesicular GABA
RT   transporter.";
RL   Nature 389:870-876(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Berks M., Smith A.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   REVISIONS.
RA   Durbin R.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Involved in the uptake of GABA into the synaptic
CC       vesicles.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Intracellular
CC       membrane vesicles.
CC   -!- SIMILARITY: Belongs to the amino acid/polyamine transporter family
CC       II.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF031935; AAB87066.1; -.
DR   EMBL; Z30423; CAA83006.2; -.
DR   PIR; S42372; S42372.
DR   PIR; T42254; T42254.
DR   WormPep; T20G5.6; CE25119.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
KW   Glycoprotein; Neurotransmitter transport; Transmembrane; Transport.
FT   DOMAIN        1     93       Cytoplasmic (Potential).
FT   TRANSMEM     94    114       Potential.
FT   DOMAIN      115    119       Vesicular lumen (Potential).
FT   TRANSMEM    120    140       Potential.
FT   DOMAIN      141    167       Cytoplasmic (Potential).
FT   TRANSMEM    168    188       Potential.
FT   DOMAIN      189    203       Vesicular lumen (Potential).
FT   TRANSMEM    204    224       Potential.
FT   DOMAIN      225    228       Cytoplasmic (Potential).
FT   TRANSMEM    229    249       Potential.
FT   DOMAIN      250    263       Vesicular lumen (Potential).
FT   TRANSMEM    264    284       Potential.
FT   DOMAIN      285    305       Cytoplasmic (Potential).
FT   TRANSMEM    306    326       Potential.
FT   DOMAIN      327    341       Vesicular lumen (Potential).
FT   TRANSMEM    342    362       Potential.
FT   DOMAIN      363    398       Cytoplasmic (Potential).
FT   TRANSMEM    399    419       Potential.
FT   DOMAIN      420    421       Vesicular lumen (Potential).
FT   TRANSMEM    422    442       Potential.
FT   DOMAIN      443    457       Cytoplasmic (Potential).
FT   TRANSMEM    458    478       Potential.
FT   DOMAIN      479    486       Vesicular lumen (Potential).
FT   CARBOHYD    337    337       N-linked (GlcNAc...) (Potential).
FT   MUTAGEN     462    462       G->R: In N2409; loss of GABA transport
FT                               into synaptic vesicles.
SQ   SEQUENCE   486 AA;  54439 MW;  6435C3B338552925 CRC64;

Query Match             31.2%;  Score 858; DB 1; Length 486;
Best Local Similarity   40.4%;  Pred. No. 3.7e-55;
Matches  180; Conservative  96; Mismatches 141; Indels  28; Gaps   8;

Qy     75  AEAPVEGDIHYQRGSGAPLPPSGSKDQVGGGGEFGGHDKPKITAWEAGWNVTNAIQGMFV 134
               :    :: :|| :        :          |:  |     :|||| ||||  |
Db     63  SEQPQKDDINKQ------------EEAKDD--GHGE----ASEPISALQAAWNVTNAIQGMFI 107

Qy    135  LGLPYAILHGGYLGLFLIIFAAVVCCYTGKILIACLYEENEDGEVVRVRDSYVAIANACC 194
               :||:|  ||||||  || ||:|||  :||:|||:    :||      |: |:||
Db    108  VGLPIAVKVGGWWSIGAMVGVAYVCYWTGVLLIECLYENG----VKKRKTYREIAD-FY 161

Qy    195  APRFPTLGGRVVNVAQIIELVMTCILYVVVSGNLMYNSPPGLPVSQKSWSIIATAVLLPC 254
               :|  |:|::| ::| |||:|||:  ||: |: |||: ::    :|:  :|:|:
Db    162  KPGF----GKWVLAAQLTELLSTCIIYLVLAADLLQSCFPS--VDKAGWMITSASLLTC 215

Qy    255  AFLKNLKAVSKFSLLCTLAHFVINILVIAYCLSRARDWAWEKVKFYIDVKKFPTSIGIIV 314
               :|| || : |:|   :||| ||||||:|: |||    :: :|  ||| :||||| ||
Db    216  SFLDDLQIVSRLSFFNAISHLIVNLIMVLYCLSFVSQWSFSTIFSLNINTLPTIVGMVV 275

Qy    315  FSYTSQIFLPSLEGNMQQPSEFHCMNWTHIAACVLKGLFALVAVLTWADETKEVITDNL 374
               ||| | ||| |||||| | :|   |||||   |:|| |:||||::  |||||||| |
Db    276  FGYTSHIFLPNLEGNMKNPAQFNVMLKWSHIAAAVFKVVFGMLGFLTFGGSLTQEEISNSL 335

Qy    375  PG-SIRAVVNIFLVAKALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGL 433
               |:  ||||| |:| ||  || |||||||| || :|| ||  :   |: :| ||:|| ||
Db    336  PNQSFKILVNLILVVKALLSYPLPFFAAVQLKNNLFGLFGYPQTPFTSCISPDKSLREWAV 395

Qy    434  TLRCALVFTLLMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLWRKLLWHQVFPD 493
               ||| | ||| |:|:| |: | |||:|| |||  |||||| ||||| ||||||||||| |
Db    396  TLRIILVLFTLFVALSVPYLVELMGLVGNITGTMLSFIWPALFHLYIKEKTLNFPEKRFD 455

Qy    494  VAIFVIGGICSVSGFVHSLEGLIEA 518
               ::| |  ||:|| ||:|| :|
Db    456  QGIIIMGCSVCISGVYFSSMELLRA 480


RESULT 13
Q8SPU0
ID   Q8SPU0          PRELIMINARY;       PRT;    164 AA.
AC   Q8SPU0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Solute carrier family 32 member 1 (Fragment).
GN   Name=SLC32A1;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Norgren R.B. Jr., Zink M.A., Jia Y., Ojeda S.R., Spindel E.R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF469843; AAL96689.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
```

DR   GO; GO:0006865; P:amino acid transport; IEA.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
FT   NON_TER   1
FT   NON_TER   164
SQ   SEQUENCE   164 AA; 18476 MW; F113174A3810A65 CRC64;

Query Match   31.0%; Score 853; DB 2; Length 164;
Best Local Similarity   100.0%; Pred. No. 2.5e-55;
Matches 164; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   208 VAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQKSWSIIATAVLLPCAFLKNLKAVSKFS 267
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 VAQIIELVMTCILYVVVSGNLMYNSFPGLPVSQKSWSIIATAVLLPCAFLKNLKAVSKFS 60

Qy   268 LLCTLAHFVINILVIAYCLSRARDWAWEKVKFYIDVKKFPISIGIIVFSYTSQIFLPSLE 327
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LLCTLAHFVINILVIAYCLSRARDWAWEKVKFYIDVKKFPISIGIIVFSYTSQIFLPSLE 120

Qy   328 GNMQQPSEFHCMNWTHIAACVLKGLFALVAYLTWADETKEVIT 371
               |||||||||||||||||||||||||||||||||||||||||||
Db   121 GNMQQPSEFHCMNWTHIAACVLKGLFALVAYLTWADETKEVIT 164

RESULT 14
Q7XUV8
ID   Q7XUV8;   PRELIMINARY;   PRT;   455 AA.
AC   Q7XUV8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   OSJNBa0072F16.7 protein.
GN   Name=OSJNBa0072F16.7;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=12447439;
RA   Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA   Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA   Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA   Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA   Wu M., Zhang R., Zhang C., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA   Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA   Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA   Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA   Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA   Han B.;
RT   "Sequence and analysis of rice chromosome 4.";
RL   Nature 420:316-320(2002).
DR   EMBL; AL606460; CAD40982.2; -.
DR   Gramene; Q7XUV8; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR   GO; GO:0006865; P:amino acid transport; IEA.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
SQ   SEQUENCE   455 AA; 47994 MW; 2F7CFE88B9CEDB00 CRC64;

Query Match   13.7%; Score 375.5; DB 2; Length 455;
Best Local Similarity   26.0%; Pred. No. 1.9e-19;
Matches 126; Conservative   84; Mismatches 200; Indels 75; Gaps 16;

Qy   48 DDLDFEHRGQLQMDILKAEGE-PCGDEGAEAPVEGDIHYQRGSGAPLPPSGSSKDQVGGGG 106
                : |: :||| : ||| :|
Db   21 EDSSTARRPDFEQPLLQAHAAVPA--RGKQEPVERDHEAQ---------CSPEADGDGA 68

Qy   107 EFGGHDKPKITAWEAGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCYTGKIL 166
                :|| :::| ||||| ||| || |||||| || |||||||| |||
Db   69 TF---------VRTCFNGLNALSGVGLLSIPYALSEGGWLSLVLLLAVAMVCCYTGLLL 118

Qy   167 IACLYEENEDGEVVRDSYVAIANACCAPRPTLGGRVVNVAQIIELVMTCILYVVVSG 226
               |:|
Db   119 RRCM-----AASPAVRGYPDIGALA-------FGAKGRLAVSAFLYAELYLVAIGFLILEG 167

Qy   227 NLMYNSFP------GLPVSQKSWSIIATA-VLLPCAFLKNLKAVSKFSLLCTLAHFVIN 278
                | :|::   :|::   ::| | : ||:| :::| |:|:|:| |:||| :|
Db   168 DNLDKLFPGTSLAVGGLVVSGKQLFVVVAVVILPTWLRSLAVLAVVS-----ASGVLA 222

Qy   279 ILVIAYCLSRARDWA--WEKVKFY-----IDVKKFPISIGIIVFSYTSQIFLPSLEGNMQ 331
               ||
Db   223 SVVVVFCVL---WAAVFDGVFHGKGRMLNVSGLPTALGLYTFCYCGHAIPTLCNSMQ 278

Qy   332 QPSEFHCMNWTHIAACVLKGLFALVAYLTWADETKEVITDNLP-GSIRAVNIF-LVAK 389
               |: :| :| :
Db   279 EKDKFSRVLVICPVACTVNYGSMAILGYLMYGDDVKSQVTLNLPEGKISSKAIYTTLIN 338

Qy   390 ALLSYPLPFFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVFTLLMAIY 449
                                                            ||:|
Db   339 PFSKYALMTPVATAIEEKLLA----------GNKRSVNVLIRTLIVVSTVVIALT 384

Qy   450 VPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLWRKLLWH-QVFFDVAIFVIGGICSVSGF 508
               | : |:| |:
Db   385 VPFFGHLMALVGSLLSVMASMLLPCICYLKIFGLTRCGRGETLLIAIIIVLGSLVAATGT 444

Qy   509 VHSLE 513
               |:
Db   445 YSSLK 449

RESULT 15
Q6YU97
ID   Q6YU97;   PRELIMINARY;   PRT;   571 AA.
AC   Q6YU97;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Putative amino acid transport protein.
GN   Name=B1170C05.10-1;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005873; BAD08181.1; -.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
SQ   SEQUENCE   571 AA; 62625 MW; 30312268725FB4DA CRC64;

Query Match   13.4%; Score 367.5; DB 2; Length 571;
Best Local Similarity   26.7%; Pred. No. 9.6e-19;
Matches 116; Conservative   83; Mismatches 175; Indels 61; Gaps 11;

Qy   113 KPKITAWE------------AGWNVTNAIQGMFVLGLPYAILHGGYLGLFLIIFAAVVCCY 161
               |: :| | | | |||||| ||: ||||||| ||||
Db   169 KPLVPAHEVPAYQQCSYTQAVMNGINVLCGVGILSTPYAIKQGGWLGLVILCLFAVLAWY 228

Qy   162 TGKILIACLYEENEDGEVVRDSYVAIANACCAPRPTLGGRVVNVAQIIELVMTCILY 221
               || | |:| ||| |: | ||
Db   229 TGVLLRRCL--DSKEG---LETYPDIGHAA----FGTTGRIAISIILYIELYACCIEY 277

Qy   222 VVVSGNLMYNSFPGLPV--------SQKSWSIIATAVLLPCAFLKNLKAVSKFSLLCTLA 273
               | |:|:|||| |:|:| :|::| | |:| |:| |:|:| |:|
Db   278 LILESDNLSKLFPNAHLTIGSMTLNSHVFFALLTTLIVMPTTWLRDLSCLSYLS-----A 332

Qy   274 HFVINILVIAYCLSRARDW------AWEKVKFYIDVKKFPISIGIIVFSYTSQIFLPSL 326
                 :   :|:| | |||: | | |: :| | |:
Db   333 GGVIASILVVVCLC---WVGVDHVGFENKGTALNLPGIPIAIGLYCYCSGHGVFPNI 388

```
Qy    327  EGNMQQPSEFHCMNWTHIAACVLKGLFALVAYLTWADETKEVITDNLPGS--IRAVVNI  384
                :::   ::   ::   |     : :: |    |   ||   :
Db    389  YSSLKNRNQFPSILFTCIGLSSILFAGAAVMGYKMFGESTESQFTLNLPENLVVSKVAVW  448

Qy    385  FLVAKALLSYPLPFAAVEVLEKSLFQEGSRAFFPACYSGDGRLKSWGLTLRCALVVFTL   444
               ||  :  :       || : |     :     ||   : :|  :| |||
Db    449  TTVANPITKYALTITPLAMSLEELLPPNQQK-----YANI-------IMLRSSLVVSTL   495

Qy    445  LMAIYVPHFALLMGLTGSLTGAGLCFLLPSLFHLRLLWRKLLWHQVFFDVAIFVIGGICS   504
             |:|: ||||||     |||| :    |::||  : ||||:    |||:  : : |::
Db    496  LIALSVPFFGLVMALVGSLLTMLVTYILPCACFLAILKRKVTWHQIAACSFIIVGVCCA   555

Qy    505  VSGFVHSLEGLIEAY   519
                 ||  |   :|:
Db    556  CVGTYSSLSKIIQNY   570
```

Search completed: November 8, 2004, 18:55:05
Job time : 473 secs